NOVEL HCV NON-STRUCTURAL POLYPEPTIDE

CROSS-REFERENCE TO RELATED APPLICATION

This application is related to provisional patent application serial no. 60/167,502, filed November 24, 1999 from which priority is claimed under 35 USC §119(e)(1) and which is incorporated herein by reference in its entirety.

FIELD OF THE INVENTION

The present invention relates to polypeptides comprising a mutant non-structural Hepatitis C virus ("HCV") polypeptide useful for immunogenic compounds for use against HCV, methods of preparing and using the same, and immunogenic compositions comprising the same. The present invention also relates to compositions comprising (a) a mutant non-structural HCV polypeptide and (b) a viral polypeptide that is not a non-structural HCV polypeptide and methods of using these compositions.

BACKGROUND OF THE INVENTION

HCV is now recognized as the major agent of chronic hepatitis and liver disease worldwide. It is estimated that HCV infects about 400 million people worldwide, corresponding to more than 3% of the world population.

Hepatitis C virus ("HCV") is a small enveloped RNA *flavivirus*, which contains a positive-stranded RNA genome of about 10 kilobases. The genome has a single uninterrupted ORF that encodes a protein of 3010-3011 amino acids. The structural proteins of HCV include a core protein (C), which is highly immunogenic, as well as two envelope proteins (E1 and E2), which likely form a heterodimer *in vivo*, and non-structural proteins NS2-NS5. It is known that the NS3 region of the virus is important for post-translational processing of the polyprotein into individual proteins, and the NS5 region encodes an RNA-dependant RNA polymerase.

Virus-specific T lymphocytes, along with neutralizing antibodies, are the mainstay of the antiviral immune defense in established viral infections. Whereas $CD8^+$ cytotoxic T cells eliminate virus-infected-cells, $CD4^+$ T helper cells are essential for the efficient regulation of the antiviral immune response. $CD4^+$ T helper cells recognize specific antigens as peptides bound to autologous HLA class II molecules (viral antigens or particles are taken up by professional antigen-presenting cells, processed to peptides, bound to HLA class II molecules in the lysosomal compartment, and transported back to the cell surface). Several observations support an important role of $CD4^+$ T cells in the elimination of HCV infection. Tsai *et al.*, 1997 Hepatology 25:449-458; Diepolder et al 1995 Lancet 346: 1—6-1009; Missale et al 1996 JCI 98: 706-714; Botarelli et al 1993; Gastro 104: 580-587; Diepolder et al 1997 J.Virol 71: 6011. Immunogenic peptides usually have a minimal length of 8-11 amino acids. However, since the peptide binding groove of HLA class II molecules seems to be open at both ends, longer peptides are tolerated. Thus peptides eluted from HLA class II molecules are typically in the range of 15-25 amino acids. HLA class II molecules are extremely polymorphic and each allele seems to have its individual requirements for peptide binding. Thus the HLA class II repertoire of a given individual determines which viral peptides can be presented to T cells. Recognition of the specific HLA-peptide complex by the T cell receptor accompanied by appropriate costimulatory signals lead to T cell activation, secretion of cytokines, and T cell proliferation.

Sub. D'

Numerous studies demonstrate that HLA Class II restricted $CD4^+$ responses are determined by stimulating peripheral blood mononuclear cells with recombinant viral antigens or peptides. Botarelli *et al.*, (1993) Gastroenterology 104:580-587; Farrari *et al.*, (1994) Hepatology 19:286-295; Minutello *et al.*, (1993) C. J. Exp. Med. 178:17-25; Hoffmann *et al.*, (1995) Hepatology 21:632-638; Iwata *et al.*, (1995) Hepatology 22:1057-1064; and Tsai.*et al.*, (1995) Hepatology 21:908-912.

Polyclonal multispecific $CD8^+$ T cell responses have been detected in patients with chronic hepatitis C. Additionally, $CD8^+$ CTL's were shown to be important in resolving acute HCV infection in chimpanzees (Cooper *et al.*, Immunity 1999). About 50% of patients with chronic hepatitis C demonstrate a detectable virus-specific $CD4^+$ T cell

response, which is most frequently directed against HCV core and/or NS4 and tends to be more common in patients who achieve sustained viral clearance during interferon-α therapy.

Depending on the pattern of lymphokines, $CD4^+$ T helper cells have been classified as TH1, TH0, or TH2. Cytokines of the TH1 type are typically IFN-γ, lymphotoxin, and interleukin-2 (IL-2), which are believed to support activation of virus-specific $CD8^+$ T cells and natural killer cells. The TH2 cytokines IL-4, IL-5, IL-10, and IL-13 are important for B cell activation and differentiation, thus inducing a humoral immune response.

During acute hepatitis C infection a strong and sustained TH1/TH0 response to NS3 and possibly to other nonstructural proteins is associated with a self-limited course of the disease. Diapolder *et al.*, (1995) Lancet 346:1006-1007, showed all $CD4^+$ T cell clones to have a TH1 or TH0 cytokine profile, suggesting that the clones support cytotoxic immune mechanisms *in vivo*. The majority of $CD4^+$ T cell clones responded to a relatively short segment of NS3, namely amino acids 1207-1278, suggesting that this region of NS3 is immunodominant for $CD4^+$ T cells. More than 70% of those who contract HCV develop chronic infection and hepatitis, and a significant portion of them progress to cirrhosis and eventually hepatocellular carcinoma. The only approved therapy at present is a 6- to 12- month course of interferon α, which leads to sustained improvement in only 20% of patients. So far, no commercial vaccine is available.

Thus, there remains a need for compositions and methods capable of promoting anti-HCV responses.

SUMMARY OF THE INVENTION

In one aspect, the present invention relates to isolated polypeptides comprising mutant hepatitis C ("HCV") polypeptides comprising at least portions of NS3, NS4, and NS5. In a preferred aspect, NS3 is encoded by a nucleic acid sequence having an N-terminal deletion to remove the catalytic domain. The NS mutant polypeptides can include NS3, NS4s, NS4b, NS5a, NS5b or portions thereof. For example, in various embodiments, the mutant NS polypeptide comprises NS3, NS4 (NS4a and NS4b) and NS5 (NS5a and NS5b). In other embodiments, the NS polypeptide consists of NS3 and NS4 (for example,

NS4a and/or NS4b) or NS3 and NS5 (for example, NS5a and/or NS5b). Other combinations of full-length or fragments of non-structural components are also contemplated.

In another preferred aspect, the polypeptides further comprise a viral polypeptide that is not a non-structural HCV polypeptide. Such polypeptides are preferably C, or antigenic fragments thereof, more preferably, truncated C of HCV. Other polypeptides are preferably E, or antigenic fragments thereof, more preferably, E1 or E2 of HCV. Such polypeptides need not be encoded by a natural HCV genome, and include, for example, truncated or otherwise mutant HCV polypeptides or polypeptides derived from other genomes, such as, for example, polypeptides of HBV. Thus, the invention includes an isolated mutant non-structural ("NS") HCV polypeptide comprising a polypeptide having a mutation in the catalytic domain of NS3 that functionally disrupts the catalytic domain. The mutation can be, for example, a deletion or a substitution mutation. In certain embodiments, the mutant NS polypeptide comprises NS3, NS4 and NS5. In other embodiments, the mutant NS polypeptides described herein further comprise a second viral polypeptide that is not NS3, NS4, or NS5 of HCV, for example an HCV Core polypeptide ("C"), or fragment thereof, or an HCV envelope protein ("E"), for example E1 and/or E2. In certain embodiments, C is truncated (*e.g.*, at amino acid 121).

In another aspect, the present invention relates to compositions comprising any of the mutant hepatitis C ("HCV") polypeptides described herein, for example polypeptides comprising at least portions of NS3, NS4, and NS5. In a preferred aspect, NS3 is encoded by a nucleic acid sequence having an N-terminal deletion to disrupt the function of the catalytic domain, for example by removing this domain. In another preferred aspect, the polypeptides further comprise a viral polypeptide that is not a non-structural HCV polypeptide. Such polypeptides are preferably C, or antigenic fragments thereof, more preferably, truncated C of HCV. Other polypeptides are preferably E, or antigenic fragments thereof, more preferably, E1 or E2 of HCV. Such polypeptides need not be encoded by a natural HCV genome, and include, for example, truncated or otherwise mutant HCV polypeptides or polypeptides derived from other genomes, such as, for example, polypeptides of HBV. In another aspect, the invention includes a composition

comprising (a) any of the polypeptides described herein; and (b) a pharmaceutically acceptable excipient (*e.g.*, carrier and/or adjuvant).

In another aspect, the invention includes an isolated and purified polynucleotide which encodes any of the mutant HCV polypeptides described herein. In certain embodiments, the invention includes a composition comprising (a) the isolated purified polynucleotide encoding any of the mutant HCV polypeptides; and (b) a pharmaceutically acceptable excipient. The polynucleotide, can be for example, DNA in a plasmid, or is in a plasmid. Additionally, the polynucleotides described herein may be included in an expression vector as shown in the attached Figures and Sequence Listings.

In another aspect, the present invention relates to host cells transformed with expression vectors comprising a nucleic acid sequence encoding a mutant HCV polypeptide comprising at least portions of NS3, NS4, and NS5. In a preferred aspect, the expression vectors of the host cells further comprises at least one nucleic acid sequence encoding a viral polypeptide that is not a non-structural HCV polypeptide. Such polypeptides are preferably C, or antigenic fragments thereof, more preferably, truncated C of HCV. Other polypeptides are preferably E, or antigenic fragments thereof, more preferably, E1 or E2 of HCV. Such polypeptides need not be encoded by a natural HCV genome, and include, for example, truncated or otherwise mutant HCV polypeptides or polypeptides derived from other genomes, such as, for example, polypeptides of HBV. In another preferred aspect the nucleic acid sequences of the expression vectors are coexpressed. In yet another preferred aspect, the host cells are yeast cells or mammalian cells.

In another aspect, the present invention relates to expression vectors comprising a nucleic acid sequence encoding a mutant HCV polypeptide comprising NS3, NS4, and NS5. In a preferred aspect, the expression vectors of the host cells further comprises at least one nucleic acid sequence encoding a viral polypeptide that is not a non-structural HCV polypeptide. Such polypeptides are preferably C, or antigenic fragments thereof, more preferably, truncated C of HCV. Other polypeptides are preferably E, or antigenic fragments thereof, more preferably, E1 or E2 of HCV. Importantly, such polypeptides need not be encoded by a natural HCV genome, such as, for example, truncated or

otherwise mutant HCV polypeptides or polypeptides derived from other genomes, such as, for example, polypeptides of HBV. In another aspect, the present invention relates to methods of preparing a mutant HCV polypeptides. In a preferred aspect, the method comprises the steps of transforming a host cell with an expression vector, said vector comprising a nucleic acid sequence encoding a mutant HCV polypeptide comprising at least portions of NS3, NS4, and NS5, and isolating said polypeptide. In another preferred aspect the HCV polypeptide further comprises a viral polypeptide that is not a non-structural HCV polypeptide. Such polypeptides are preferably C, or antigenic fragments thereof, more preferably, truncated C of HCV. Other polypeptides are preferably E, or antigenic fragments thereof, more preferably, E1or E2 of HCV. Such polypeptides need not be encoded by a natural HCV genome, and include, for example, truncated or otherwise mutant HCV polypeptides or polypeptides derived from other genomes, such as, for example, polypeptides of HBV. In another preferred aspect the host cells are yeast cells or mammalian cells.

In another aspect, the present invention relates to antibodies which specifically bind to mutant HCV polypeptide comprising NS3, NS4, and NS5, and to methods of making and using the same. In a preferred aspect, the HCV polypeptide further comprises a viral polypeptide that is not a non-structural HCV polypeptide. Such polypeptides are preferably C, or antigenic fragments thereof, more preferably, truncated C of HCV. Other polypeptides are preferably E, or antigenic fragments thereof, more preferably, E1 or E2 of HCV. Such polypeptides need not be encoded by a natural HCV genome, such as, for example, truncated or otherwise mutant HCV polypeptides or polypeptides derived from other genomes, and include, for example, polypeptides of HBV. In another preferred aspect, the antibody is either monoclonal or polyclonal.

In yet another aspect, a method of preparing a mutant NS HCV polypeptide, wherein the method comprises the steps of (a) transforming a host cell with any of the expression vectors described herein, under conditions wherein the polypeptide is expressed; and (b) isolating the polypeptide. The host cell can be, for example, a yeast cell, a mammalian cell a plant cell or an insect cell. The polypeptide can be expressed and isolated intracellularly or can be secreted and isolated from the surrounding environment.

In a still further aspect, a method of eliciting an immune response in a subject is provided. The immune response can be elicited by administering any of the polynucleotides and/or polypeptides described herein in one or multiple doses.

These and other embodiments of the subject invention will readily occur to those of skill in the art in light of the disclosure herein.

BRIEF DESCRIPTION OF THE FIGURES

FIG. 1 shows the cloning scheme for generating pCMV-NS35.

FIG. 2 shows the 9621bp vector pCMV-NS35.

FIG. 3 shows the nucleic acid sequence of pCMV-NS35 (SEQ ID NO:1), including the nucleic acid sequence of the NS35 ORF, and also the translation of NS35 (SEQ ID NO:2).

FIG. 4 shows the 9621bp pCMV-delNS35.

FIG. 5 shows the nucleic acid sequence of pCMV-delNS35 (SEQ ID NO:3), including the nucleic acid sequence of the delNS35 ORF, and also the translation of the delNS35 polypeptide (SEQ ID NO:4).

FIG. 6 shows the 4276bp pCMV-II.

FIG. 7 shows the nucleic acid sequence of pCMV-II (SEQ ID NO:5).

FIG. 8 shows the 6300bp pCMV-NS34A.

FIG. 9 shows the nucleic acid sequence of pCMV-NS34A (SEQ ID NO:6), including the nucleic acid sequence of the NS34A ORF, and also the translation of NS34A (SEQ ID NO:7).

FIG. 10 shows the cloning scheme for generating pd.ΔNS3NS5.

FIG. 11 shows the nucleic and amino acid sequences of pd.ΔNS3NS5 (SEQ ID NO:8 and 9).

FIG. 12 shows the Western blot of proteins expressed by S. cerevisiae strain AD3 transformed with pd.ΔNS3NS5.

FIG. 13 shows the cloning scheme for generating pd.ΔNS3NS5.pj.

FIG. 14 shows the nucleic and amino acid sequences of pd.ΔNS3NS5.pj (SEQ ID NO:10 and 11).

FIG. 15 shows the Western blot of proteins expressed by S. cerevisiae strain AD3 transformed with pd.ΔNS3NS5.pj, specifically demonstrating the expression of ΔNS3NS5 polypeptide.

FIG. 16 shows the cloning scheme for generating pdΔNS3NS5.pj.core121RT and pdΔNS3NS5.pj.core173RT.

FIG. 17 shows the nucleic and amino acid sequences of pd.ΔNS3NS5.pj.core121 (SEQ ID NO:12 and 13).

FIG. 18 shows the nucleic and amino acid sequences of pd.ΔNS3NS5.pj.core173 (SEQ ID NO:14 and 15).

FIG. 19 shows the Western blot of proteins expressed by S. cerevisiae strain AD3 transformed with pd.ΔNS3NS5.pj, specifically demonstrating the expression of ΔNS3NS5.core121 and ΔNS3NS5.core173 polypeptides. Lanes 1 and 7 show See Blue Standards. Lane 2 shows control yeast plasmid. Lanes 3 and 4 show ΔNS3NS5.core121RT polypeptide, colonies 1 and 2. Lanes 5 and 6 show ΔNS3NS5.core173RT polypeptide, colonies 3 and 4.

FIG. 20 shows the cloning scheme for generating pdΔNS3NS5.pj.core140RT and pdΔNS3NS5.pj.core150RT.

FIG. 21 shows the nucleic and amino acid sequences of pd.ΔNS3NS5.pj.core140 (SEQ ID NO:16 and 17).

FIG. 22 shows the nucleic and amino acid sequences of pd.ΔNS3NS5.pj.core150 (SEQ ID NO:18 and 19).

FIG. 23 shows the Western blot of proteins expressed by S. cerevisiae strain AD3 transformed with pd.ΔNS3NS5.pj, specifically demonstrating the expression of ΔNS3NS5core140 and ΔNS3NS5core150 polypeptides. Lane 1 shows See Blue Standards. Lanes 2 and 3 show ΔNS3NS55core140RT polypeptide, colonies 5 and 6. Lanes 4 and 5 show ΔNS3NS5core150RT polypeptide, colonies 7 and 8. Lane 6 shows control yeast plasmid. Lane 7 shows ΔNS3NS5core121RT polypeptide, colony 1. Lane 8 shows ΔNS3NS5core173RT polypeptide, colony 5.

DETAILED DESCRIPTION OF THE INVENTION

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology, microbiology, recombinant DNA techniques, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature. See e.g., Sambrook, et al., MOLECULAR CLONING; A LABORATORY MANUAL (1989); DNA CLONING, VOLUMES I AND II (D. N. Glover ed. 1985); OLIGONUCLEOTIDE SYNTHESIS (M. J. Gait ed., 1984); NUCLEIC ACID HYBRIDIZATION (B. D. Hames & S. J. Higgins eds. 1984); TRANSCRIPTION AND TRANSLATION (B. D. Hames & S. J. Higgins eds. 1984); ANIMAL CELL CULTURE (R. I. Freshney ed. 1986); IMMOBILIZED CELLS AND ENZYMES (IRL Press, 1986); B. Perbal, A PRACTICAL GUIDE TO MOLECULAR CLONING (1984); the series, METHODS OF ENZYMOLOGY (Academic Press, Inc.); GENE TRANSFER VECTORS FOR MAMMALIAN CELLS (J. H. Miller and M. P. Calos eds. 1987, Cold Springs Harbor Laboratory), Methods in Enzymology Vol. 154 and Vol. 155 (Wu and Grossman, and Wu, eds., respectively); Mayer and Walker eds. (1987), IMMUNOHISTOCHEMICAL METHODS IN CELL AND MOLECULAR BIOLOGY (Academic Press, London); Scopes, (1987), PROTEIN PURIFICATION: PRINCIPALS AND PRACTICE, Second Edition (Springer-Verlag, New York); and HANDBOOK OF EXPERIMENTAL IMMUNOLOGY, VOLUMES I-IV (D. M. Weir and C. C. Blackwell eds. 1986).

All publications, patents and patent applications cited herein, whether *supra* or *infra*, are hereby incorporated by reference in their entirety.

It must be noted that, as used in this specification and the appended claims, the singular forms "a", "an" and "the" include plural referents unless the content clearly dictates otherwise. Thus, for example, reference to "an antigen" includes a mixture of two or more antigens, and the like.

I. Definitions

In describing the present invention, the following terms will be employed, and are intended to be defined as indicated below.

The term "hepatitis C virus" (HCV) refers to an agent causative of Non-A, Non-B Hepatitis (NANBH). The nucleic acid sequence and putative amino acid sequence of HCV is described in U.S. Patent Nos. 5,856,437 and 5,350,671. The disease caused by HCV is called hepatitis C, formerly called NANBH. The term HCV, as used herein, denotes a viral species of which pathenogenic strains cause NANBH, as well as attenuated strains or defective interfering particles derived therefrom.

HCV is a member of the viral family flaviviridae. The morphology and composition of Flavivirus particles are known, and are discussed in Reed et al., *Curr. Stud. Hematol. Blood Transfus.* (1998), 62:1-37; HEPATITIS C VIRUSES IN FIELDS VIROLOGY (B.N. Fields, D.M. Knipe, P.M. Howley, eds.) (3d ed. 1996). It has recently been found that portions of the HCV genome are also homologous to pestiviruses. Generally, with respect to morphology, Flaviviruses contain a central nucleocapsid surrounded by a lipid bilayer. Virions are spherical and have a diameter of about 40-50 nm. Their cores are about 25-30 nm in diameter. Along the outer surface of the virion envelope are projections that are about 5-10 nm long with terminal knobs about 2 nm in diameter.

The HCV genome is comprised of RNA. It is known that RNA containing viruses have relatively high rates of spontaneous mutation. Therefore, there can be multiple strains, which can be virulent or avirulent, within the HCV class or species. The ORF of HCV, including the translation spans of the core, non-structural, and envelope proteins, is shown in U.S. Patent Nos. 5,856,437 and 5,350,671.

The terms "polypeptide" and "protein" refer to a polymer of amino acid residues and are not limited to a minimum length of the product. Thus, peptides, oligopeptides, dimers, multimers, and the like, are included within the definition. Both full-length proteins and fragments thereof are encompassed by the definition. The terms also include postexpression modifications of the polypeptide, for example, glycosylation, acetylation, phosphorylation and the like. Furthermore, for purposes of the present invention, a

"polypeptide" refers to a protein which includes modifications, such as deletions, additions and substitutions (generally conservative in nature), to the native sequence, so long as the protein maintains the desired activity. These modifications may be deliberate, as through site-directed mutagenesis, or may be accidental, such as through mutations of hosts which produce the proteins or errors due to PCR amplification.

An HCV polypeptide is a polypeptide, as defined above, derived from the HCV polyprotein. The polypeptide need not be physically derived from HCV, but may be synthetically or recombinantly produced. Moreover, the polypeptide may be derived from any of the various HCV strains, such as from strains 1, 2, 3 or 4 of HCV. A number of conserved and variable regions are known between these strains and, in general, the amino acid sequences of epitopes derived from these regions will have a high degree of sequence homology, e.g., amino acid sequence homology of more than 30%, preferably more than 40%, when the two sequences are aligned and homology determined by any of the programs or algorithms described herein. Thus, for example, the term "NS4" polypeptide refers to native NS4 from any of the various HCV strains, as well as NS4 analogs, muteins and immunogenic fragments, as defined further below.

Further, the terms "ΔNS35," "delNS35," "ΔNS3NS5," and "ΔNS3-5" as used herein refer to a mutant polypeptide, comprising at least portions of NS3, NS4, or NS5, comprising a deletion in, or mutation of, the NS3 protease active site region to render the protease non-functional. In one embodiment, ΔNS3-5 comprises amino acids 1242-3011, as shown in FIG. 5, or polypeptides substantially homologous thereto. It will be readily apparent to one of ordinary skill in the art how to determine that NS3 protease has been rendered non-functional. If the protease is functional, one will obtain protein of the expected molecular weight upon expression. As set forth in Example 2 and Figure 15, using SDS-page, 4-20%, a protein having a molecular weight of approximately 194kD was obtained when strain AD3 was transformed with pd.ΔNS3NS5.PJ clone #5. One skilled in the art could readily determine whether a protein of the desired molecular weight was expressed for any given deletion or mutation.

The terms "analog" and "mutein" refer to biologically active derivatives of the reference molecule, or fragments of such derivatives, that retain desired activity, such as

the ability to stimulate a cell-mediated immune response, as defined below. In general, the term "analog" refers to compounds having a native polypeptide sequence and structure with one or more amino acid additions, substitutions (generally conservative in nature) and/or deletions, relative to the native molecule, so long as the modifications do not destroy immunogenic activity. The term "mutein" refers to peptides having one or more peptide mimics ("peptoids"), such as those described in International Publication No. WO 91/04282. Preferably, the analog or mutein has at least the same immunoactivity as the native molecule. Methods for making polypeptide analogs and muteins are known in the art and are described further below.

Particularly preferred analogs include substitutions that are conservative in nature, i.e., those substitutions that take place within a family of amino acids that are related in their side chains. Specifically, amino acids are generally divided into four families: (1) acidic -- aspartate and glutamate; (2) basic -- lysine, arginine, histidine; (3) non-polar -- alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan; and (4) uncharged polar -- glycine, asparagine, glutamine, cysteine, serine threonine, tyrosine. Phenylalanine, tryptophan, and tyrosine are sometimes classified as aromatic amino acids. For example, it is reasonably predictable that an isolated replacement of leucine with isoleucine or valine, an aspartate with a glutamate, a threonine with a serine, or a similar conservative replacement of an amino acid with a structurally related amino acid, will not have a major effect on the biological activity. For example, the polypeptide of interest may include up to about 5-10 conservative or non-conservative amino acid substitutions, or even up to about 15-25 conservative or non-conservative amino acid substitutions, or any integer between 5-25, so long as the desired function of the molecule remains intact. One of skill in the art may readily determine regions of the molecule of interest that can tolerate change by reference to Hopp/Woods and Kyte-Doolittle plots, well known in the art.

By "fragment" is intended a polypeptide consisting of only a part of the intact full-length polypeptide sequence and structure. The fragment can include a C-terminal deletion and/or an N-terminal deletion of the native polypeptide. An "immunogenic fragment" of a particular HCV protein will generally include at least about 5-10 contiguous amino acid residues of the full-length molecule, preferably at least about 15-25 contiguous amino acid

residues of the full-length molecule, and most preferably at least about 20-50 or more contiguous amino acid residues of the full-length molecule, that define an epitope, or any integer between 5 amino acids and the full-length sequence, provided that the fragment in question retains immunogenic activity, as measured by the assays described herein. For a description of various HCV epitopes, see, e.g., Chien et al., *Proc. Natl. Acad. Sci. USA* (1992) 89:10011-10015; Chien et al., *J. Gastroent. Hepatol.* (1993) 8:S33-39; Chien et al., International Publication No. WO 93/00365; Chien, D.Y., International Publication No. WO 94/01778; commonly owned, allowed U.S. Patent Application Serial Nos. 08/403,590 and 08/444,818.

The term "epitope" as used herein refers to a sequence of at least about 3 to 5, preferably about 5 to 10 or 15, and not more than about 1,000 amino acids (or any integer therebetween), which define a sequence that by itself or as part of a larger sequence, binds to an antibody generated in response to such sequence. There is no critical upper limit to the length of the fragment, which may comprise nearly the full-length of the protein sequence, or even a fusion protein comprising two or more epitopes from the HCV polyprotein. An epitope for use in the subject invention is not limited to a polypeptide having the exact sequence of the portion of the parent protein from which it is derived. Indeed, viral genomes are in a state of constant flux and contain several variable domains which exhibit relatively high degrees of variability between isolates. Thus the term "epitope" encompasses sequences identical to the native sequence, as well as modifications to the native sequence, such as deletions, additions and substitutions (generally conservative in nature).

Regions of a given polypeptide that include an epitope can be identified using any number of epitope mapping techniques, well known in the art. See, e.g., *Epitope Mapping Protocols* in Methods in Molecular Biology, Vol. 66 (Glenn E. Morris, Ed., 1996) Humana Press, Totowa, New Jersey. For example, linear epitopes may be determined by e.g., concurrently synthesizing large numbers of peptides on solid supports, the peptides corresponding to portions of the protein molecule, and reacting the peptides with antibodies while the peptides are still attached to the supports. Such techniques are known in the art and described in, e.g., U.S. Patent No. 4,708,871; Geysen et al. (1984) *Proc.*

Natl. Acad. Sci. USA 81:3998-4002; Geysen et al. (1986) *Molec. Immunol.* 23:709-715, all incorporated herein by reference in their entireties. Similarly, conformational epitopes are readily identified by determining spatial conformation of amino acids such as by, e.g., x-ray crystallography and 2-dimensional nuclear magnetic resonance. See, e.g., *Epitope Mapping Protocols*, *supra*. Antigenic regions of proteins can also be identified using standard antigenicity and hydropathy plots, such as those calculated using, e.g., the Omiga version 1.0 software program available from the Oxford Molecular Group. This computer program employs the Hopp/Woods method, Hopp et al., *Proc. Natl. Acad. Sci USA* (1981) 78:3824-3828 for determining antigenicity profiles, and the Kyte-Doolittle technique, Kyte et al., *J. Mol. Biol.* (1982) 157:105-132 for hydropathy plots.

As used herein, the term "conformational epitope" refers to a portion of a full-length protein, or an analog or mutein thereof, having structural features native to the amino acid sequence encoding the epitope within the full-length natural protein. Native structural features include, but are not limited to, glycosylation and three dimensional structure. Preferably, a conformational epitope is produced recombinantly and is expressed in a cell from which it is extractable under conditions which preserve its desired structural features, e.g. without denaturation of the epitope. Such cells include bacteria, yeast, insect, and mammalian cells. Expression and isolation of recombinant conformational epitopes from the HCV polyprotein are described in e.g., International Publication Nos. WO 96/04301, WO 94/01778, WO 95/33053, WO 92/08734, which applications are herein incorporated by reference in their entirety.

An "immunological response" to an HCV antigen (including both polypeptide and polynucleotides encoding polypeptides that are expressed *in vivo*) or composition is the development in a subject of a humoral and/or a cellular immune response to molecules present in the composition of interest. For purposes of the present invention, a "humoral immune response" refers to an immune response mediated by antibody molecules, while a "cellular immune response" is one mediated by T-lymphocytes and/or other white blood cells. One important aspect of cellular immunity involves an antigen-specific response by cytolytic T-cells ("CTLs"). CTLs have specificity for peptide antigens that are presented in association with proteins encoded by the major histocompatibility complex (MHC) and

expressed on the surfaces of cells. CTLs help induce and promote the intracellular destruction of intracellular microbes, or the lysis of cells infected with such microbes. Another aspect of cellular immunity involves an antigen-specific response by helper T-cells. Helper T-cells act to help stimulate the function, and focus the activity of, nonspecific effector cells against cells displaying peptide antigens in association with MHC molecules on their surface. A "cellular immune response" also refers to the production of cytokines, chemokines and other such molecules produced by activated T-cells and/or other white blood cells, including those derived from CD4+ and CD8+ T-cells.

A composition or vaccine that elicits a cellular immune response may serve to sensitize a vertebrate subject by the presentation of antigen in association with MHC molecules at the cell surface. The cell-mediated immune response is directed at, or near, cells presenting antigen at their surface. In addition, antigen-specific T-lymphocytes can be generated to allow for the future protection of an immunized host.

The ability of a particular antigen to stimulate a cell-mediated immunological response may be determined by a number of assays, such as by lymphoproliferation (lymphocyte activation) assays, CTL cytotoxic cell assays, or by assaying for T-lymphocytes specific for the antigen in a sensitized subject. Such assays are well known in the art. See, e.g., Erickson et al., *J. Immunol.* (1993) 151:4189-4199; Doe et al., *Eur. J. Immunol.* (1994) 24:2369-2376; and the examples below.

Thus, an immunological response as used herein may be one which stimulates the production of CTLs, and/or the production or activation of helper T- cells. The antigen of interest may also elicit an antibody-mediated immune response. Hence, an immunological response may include one or more of the following effects: the production of antibodies by B-cells; and/or the activation of suppressor T-cells and/or γδ T-cells directed specifically to an antigen or antigens present in the composition or vaccine of interest. These responses may serve to neutralize infectivity, and/or mediate antibody-complement, or antibody dependent cell cytotoxicity (ADCC) to provide protection or alleviation of symptoms to an immunized host. Such responses can be determined using standard immunoassays and neutralization assays, well known in the art.

A "coding sequence" or a sequence which "encodes" a selected polypeptide, is a nucleic acid molecule which is transcribed (in the case of DNA) and translated (in the case of mRNA) into a polypeptide *in vitro* or *in vivo* when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxy) terminus. A transcription termination sequence may be located 3' to the coding sequence.

A "nucleic acid" molecule or "polynucleotide" can include both double- and single-stranded sequences and refers to, but is not limited to, cDNA from viral, procaryotic or eucaryotic mRNA, genomic DNA sequences from viral (e.g. DNA viruses and retroviruses) or procaryotic DNA, and especially synthetic DNA sequences. The term also captures sequences that include any of the known base analogs of DNA and RNA.

"Operably linked" refers to an arrangement of elements wherein the components so described are configured so as to perform their desired function. Thus, a given promoter operably linked to a coding sequence is capable of effecting the expression of the coding sequence when the proper transcription factors, etc., are present. The promoter need not be contiguous with the coding sequence, so long as it functions to direct the expression thereof. Thus, for example, intervening untranslated yet transcribed sequences can be present between the promoter sequence and the coding sequence, as can transcribed introns, and the promoter sequence can still be considered "operably linked" to the coding sequence.

"Recombinant" as used herein to describe a nucleic acid molecule means a polynucleotide of genomic, cDNA, viral, semisynthetic, or synthetic origin which, by virtue of its origin or manipulation is not associated with all or a portion of the polynucleotide with which it is associated in nature. The term "recombinant" as used with respect to a protein or polypeptide means a polypeptide produced by expression of a recombinant polynucleotide. In general, the gene of interest is cloned and then expressed in transformed organisms, as described further below. The host organism expresses the foreign gene to produce the protein under expression conditions.

A "control element" refers to a polynucleotide sequence which aids in the expression of a coding sequence to which it is linked. The term includes promoters,

transcription termination sequences, upstream regulatory domains, polyadenylation signals, untranslated regions, including 5'-UTRs and 3'-UTRs and when appropriate, leader sequences and enhancers, which collectively provide for the transcription and translation of a coding sequence in a host cell.

A "promoter" as used herein is a DNA regulatory region capable of binding RNA polymerase in a host cell and initiating transcription of a downstream (3' direction) coding sequence operably linked thereto. For purposes of the present invention, a promoter sequence includes the minimum number of bases or elements necessary to initiate transcription of a gene of interest at levels detectable above background. Within the promoter sequence is a transcription initiation site, as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase. Eucaryotic promoters will often, but not always, contain "TATA" boxes and "CAT" boxes.

A control sequence "directs the transcription" of a coding sequence in a cell when RNA polymerase will bind the promoter sequence and transcribe the coding sequence into mRNA, which is then translated into the polypeptide encoded by the coding sequence.

"Expression cassette" or "expression construct" refers to an assembly which is capable of directing the expression of the sequence(s) or gene(s) of interest. The expression cassette includes control elements, as described above, such as a promoter which is operably linked to (so as to direct transcription of) the sequence(s) or gene(s) of interest, and often includes a polyadenylation sequence as well. Within certain embodiments of the invention, the expression cassette described herein may be contained within a plasmid construct. In addition to the components of the expression cassette, the plasmid construct may also include, one or more selectable markers, a signal which allows the plasmid construct to exist as single-stranded DNA (e.g., a M13 origin of replication), at least one multiple cloning site, and a "mammalian" origin of replication (e.g., a SV40 or adenovirus origin of replication).

"Transformation," as used herein, refers to the insertion of an exogenous polynucleotide into a host cell, irrespective of the method used for insertion: for example, transformation by direct uptake, transfection, infection, and the like. For particular methods of transfection, see further below. The exogenous polynucleotide may be

maintained as a nonintegrated vector, for example, an episome, or alternatively, may be integrated into the host genome.

A "host cell" is a cell which has been transformed, or is capable of transformation, by an exogenous DNA sequence.

By "isolated" is meant, when referring to a polypeptide, that the indicated molecule is separate and discrete from the whole organism with which the molecule is found in nature or is present in the substantial absence of other biological macromolecules of the same type. The term "isolated" with respect to a polynucleotide is a nucleic acid molecule devoid, in whole or part, of sequences normally associated with it in nature; or a sequence, as it exists in nature, but having heterologous sequences in association therewith; or a molecule disassociated from the chromosome.

The term "purified" as used herein preferably means at least 75% by weight, more preferably at least 85% by weight, more preferably still at least 95% by weight, and most preferably at least 98% by weight, of biological macromolecules of the same type are present.

"Homology" refers to the percent identity between two polynucleotide or two polypeptide moieties. Two DNA, or two polypeptide sequences are "substantially homologous" to each other when the sequences exhibit at least about 50% , preferably at least about 75%, more preferably at least about 80%-85%, preferably at least about 90%, and most preferably at least about 95%-98%, or more, sequence identity over a defined length of the molecules. As used herein, substantially homologous also refers to sequences showing complete identity to the specified DNA or polypeptide sequence. The term "substantially homologous" as used herein in reference to ΔNS35 generally refers to an HCV nucleic or amino acid sequence that is at least 60% identical to the entire sequence of the polypeptide encoded by ΔNS35 (see FIG. 5), where the sequence identity is preferably at least 75%, more preferably at least 80%, still more preferably at least about 85%, especially more than about 90%, most preferably 95% or greater, particularly 98% or greater. These homologous polypeptides include fragments, including mutants and allelic variants of the fragments. Identity between the two sequences is preferably determined by the Smith-Waterman homology search algorithm as implemented in the MPSRCH program

(Oxford Molecular), using an affine gap search with parameters *gap open penalty*=12 and *gap extension penalty*=1. Thus, for example, the present invention includes an isolate which is 80% identical to a polypeptide encoded by ΔNS35. In some aspects of the invention, the polypeptide of the present invention is substantially homologous to the ΔNS35.

In general, "identity" refers to an exact nucleotide-to-nucleotide or amino acid-to-amino acid correspondence of two polynucleotides or polypeptide sequences, respectively. Percent identity can be determined by a direct comparison of the sequence information between two molecules by aligning the sequences, counting the exact number of matches between the two aligned sequences, dividing by the length of the shorter sequence, and multiplying the result by 100. Readily available computer programs can be used to aid in the analysis, such as ALIGN, Dayhoff, M.O. in *Atlas of Protein Sequence and Structure* M.O. Dayhoff ed., 5 Suppl. 3:353-358, National biomedical Research Foundation, Washington, DC, which adapts the local homology algorithm of Smith and Waterman *Advances in Appl. Math.* 2:482-489, 1981 for peptide analysis. Programs for determining nucleotide sequence identity are available in the Wisconsin Sequence Analysis Package, Version 8 (available from Genetics Computer Group, Madison, WI) for example, the BESTFIT, FASTA and GAP programs, which also rely on the Smith and Waterman algorithm. These programs are readily utilized with the default parameters recommended by the manufacturer and described in the Wisconsin Sequence Analysis Package referred to above. For example, percent identity of a particular nucleotide sequence to a reference sequence can be determined using the homology algorithm of Smith and Waterman with a default scoring table and a gap penalty of six nucleotide positions.

Sub. D2

Another method of establishing percent identity in the context of the present invention is to use the MPSRCH package of programs copyrighted by the University of Edinburgh, developed by John F. Collins and Shane S. Sturrok, and distributed by IntelliGenetics, Inc. (Mountain View, CA). From this suite of packages the Smith-Waterman algorithm can be employed where default parameters are used for the scoring table (for example, gap open penalty of 12, gap extension penalty of one, and a gap of six). From the data generated the "Match" value reflects "sequence identity." Other suitable

Sub D2

programs for calculating the percent identity or similarity between sequences are generally known in the art, for example, another alignment program is BLAST, used with default parameters. For example, BLASTN and BLASTP can be used using the following default parameters: genetic code = standard; filter = none; strand = both; cutoff = 60; expect = 10; Matrix = BLOSUM62; Descriptions = 50 sequences; sort by = HIGH SCORE; Databases = non-redundant, GenBank + EMBL + DDBJ + PDB + GenBank CDS translations + Swiss protein + Spupdate + PIR. Details of these programs can be found at the following internet address: http://www.ncbi.nlm.gov/cgi-bin/BLAST.

Alternatively, homology can be determined by hybridization of polynucleotides under conditions which form stable duplexes between homologous regions, followed by digestion with single-stranded-specific nuclease(s), and size determination of the digested fragments. DNA sequences that are substantially homologous can be identified in a Southern hybridization experiment under, for example, stringent conditions, as defined for that particular system. Defining appropriate hybridization conditions is within the skill of the art. See, e.g., Sambrook et al., *supra*; *DNA Cloning*, *supra*; *Nucleic Acid Hybridization*, *supra*.

"Stringency" refers to conditions in a hybridization reaction that favor association of very similar sequences over sequences that differ. For example, the combination of temperature and salt concentration should be chosen that is approximately 120 to 200°C below the calculated Tm of the hybrid under study. The temperature and salt conditions can often be determined empirically in preliminary experiments in which samples of genomic DNA immobilized on filters are hybridized to the sequence of interest and then washed under conditions of different stringencies. See Sambrook *et al.* at page 9.50.

Variables to consider when performing, for example, a Southern blot are (1) the complexity of the DNA being blotted and (2) the homology between the probe and the sequences being detected. The total amount of the fragment(s) to be studied can vary a magnitude of 10, from 0.1 to 1µg for a plasmid or phage digest to 10^{-9} to 10^{-8} g for a single copy gene in a highly complex eukaryotic genome. For lower complexity polynucleotides, substantially shorter blotting, hybridization, and exposure times, a smaller amount of starting polynucleotides, and lower specific activity of probes can be used. For example, a

single-copy yeast gene can be detected with an exposure time of only 1 hour starting with 1 μg of yeast DNA, blotting for two hours, and hybridizing for 4-8 hours with a probe of 10^8 cpm/μg. For a single-copy mammalian gene a conservative approach would start with 10 μg of DNA, blot overnight, and hybridize overnight in the presence of 10% dextran sulfate using a probe of greater than 10^8 cpm/μg, resulting in an exposure time of ~24 hours.

Several factors can affect the melting temperature (Tm) of a DNA-DNA hybrid between the probe and the fragment of interest, and consequently, the appropriate conditions for hybridization and washing. In many cases the probe is not 100% homologous to the fragment. Other commonly encountered variables include the length and total G+C content of the hybridizing sequences and the ionic strength and formamide content of the hybridization buffer. The effects of all of these factors can be approximated by a single equation:

$$Tm = 81 + 16.6(\log_{10}Ci) + 0.4[\%(G + C)] - 0.6(\%formamide) - 600/n - 1.5(\%mismatch).$$

where Ci is the salt concentration (monovalent ions) and *n* is the length of the hybrid in base pairs (slightly modified from Meinkoth & Wahl (1984) *Anal. Biochem.* 138: 267-284). In general, convenient hybridization temperatures in the presence of 50% formamide are 42°C for a probe with is 95% to 100% homologous to the target fragment, 37°C for 90% to 95% homology, and 32°C for 85% to 90% homology. For lower homologies, formamide content should be lowered and temperature adjusted accordingly, using the equation above. If the homology between the probe and the target fragment are not known, the simplest approach is to start with both hybridization and wash conditions which are nonstringent. If non-specific bands or high background are observed after autoradiography, the filter can be washed at high stringency and reexposed. If the time required for exposure makes this approach impractical, several hybridization and/or washing stringencies should be tested in parallel.

By "nucleic acid immunization" is meant the introduction of a nucleic acid molecule encoding one or more selected antigens into a host cell, for the *in vivo* expression of the antigen or antigens. The nucleic acid molecule can be introduced directly into the recipient subject, such as by injection, inhalation, oral, intranasal and mucosal administration, or the like, or can be introduced *ex vivo*, into cells which have been

removed from the host. In the latter case, the transformed cells are reintroduced into the subject where an immune response can be mounted against the antigen encoded by the nucleic acid molecule.

An "open reading frame" or ORF is a region of a polynucleotide sequence which encodes a polypeptide; this region can represent a portion of a coding sequence or a total coding sequence.

As used herein, the term "antibody" refers to a polypeptide or group of polypeptides which comprise at least one antigen binding site. An "antigen binding site" is formed from the folding of the variable domains of an antibody molecule(s) to form three-dimensional binding sites with an internal surface shape and charge distribution complementary to the features of an epitope of an antigen, which allows specific binding to form an antibody-antigen complex. An antigen binding site may be formed from a heavy- and/or light-chain domain (VH and VL, respectively), which form hypervariable loops which contribute to antigen binding. The term "antibody" includes, without limitation, polyclonal antibodies, monoclonal antibodies, chimeric antibodies, altered antibodies, univalent antibodies, Fab proteins, and single-domain antibodies. In many cases, the binding phenomena of antibodies to antigens is equivalent to other ligand/anti-ligand binding.

If polyclonal antibodies are desired, a selected mammal (e.g., mouse, rabbit, goat, horse, etc.) is immunized with an immunogenic polypeptide bearing an HCV epitope(s). Serum from the immunized animal is collected and treated according to known procedures. If serum containing polyclonal antibodies to an HCV epitope contains antibodies to other antigens, the polyclonal antibodies can be purified by immunoaffinity chromatography. Techniques for producing and processing polyclonal antisera are known in the art, see for example, Mayer and Walker, eds. (1987) IMMUNOCHEMICAL METHODS IN CELL AND MOLECULAR BIOLOGY (Academic Press, London).

Monoclonal antibodies directed against HCV epitopes can also be readily produced by one skilled in the art. The general methodology for making monoclonal antibodies by hybridomas is well known. Immortal antibody-producing cell lines can be created by cell fusion, and also by other techniques such as direct transformation of B lymphocytes with oncogenic DNA, or transfection with Epstein-Barr virus. See, e.g., M. Schreier et al.

(1980) HYBRIDOMA TECHNIQUES; Hammerling et al. (1981), MONOCLONAL ANTIBODIES AND T-CELL HYBRIDOMAS; Kennett et al. (1980) MONOCLONAL ANTIBODIES; see also, U.S. Pat. Nos. 4,341,761; 4,399,121; 4,427,783; 4,444,887; 4,466,917; 4,472,500; 4,491,632; and 4,493,890. Panels of monoclonal antibodies produced against HCV epitopes can be screened for various properties; i.e., for isotype, epitope affinity, etc. As used herein, a "single domain antibody" (dAb) is an antibody which is comprised of an HL domain, which binds specifically with a designated antigen. A dAb does not contain a VL domain, but may contain other antigen binding domains known to exist to antibodies, for example, the kappa and lambda domains. Methods for preparing dabs are known in the art. See, for example, Ward et al, Nature 341: 544 (1989).

Antibodies can also be comprised of VH and VL domains, as well as other known antigen binding domains. Examples of these types of antibodies and methods for their preparation and known in the art (see, e.g., U.S. Pat. No. 4,816,467, which is incorporated herein by reference), and include the following. For example, "vertebrate antibodies" refers to antibodies which are tetramers or aggregates thereof, comprising light and heavy chains which are usually aggregated in a "Y" configuration and which may or may not have covalent linkages between the chains. In vertebrate antibodies, the amino acid sequences of the chains are homologous with those sequences found in antibodies produced in vertebrates, whether in situ or in vitro (for example, in hybridomas). Vertebrate antibodies include, for example, purified polyclonal antibodies and monoclonal antibodies, methods for the preparation of which are described infra.

"Hybrid antibodies" are antibodies where chains are separately homologous with reference to mammalian antibody chains and represent novel assemblies of them, so that two different antigens are precipitable by the tetramer or aggregate. In hybrid antibodies, one pair of heavy and light chains are homologous to those found in an antibody raised against a first antigen, while a second pair of chains are homologous to those found in an antibody raised against a second antibody. This results in the property of "divalence", i.e., the ability to bind two antigens simultaneously. Such hybrids can also be formed using chimeric chains, as set forth below.

"Chimeric antibodies" refers to antibodies in which the heavy and/or light chains are fusion proteins. Typically, one portion of the amino acid sequences of the chain is homologous to corresponding sequences in an antibody derived from a particular species or a particular class, while the remaining segment of the chain is homologous to the sequences derived from another species and/or class. Usually, the variable region of both light and heavy chains mimics the variable regions or antibodies derived from one species of vertebrates, while the constant portions are homologous to the sequences in the antibodies derived from another species of vertebrates. However, the definition is not limited to this particular example. Also included is any antibody in which either or both of the heavy or light chains are composed of combinations of sequences mimicking the sequences in antibodies of different sources, whether these sources be from differing classes or different species of origin, and whether or not the fusion point is at the variable/constant boundary. Thus, it is possible to produce antibodies in which neither the constant nor the variable region mimic know antibody sequences. It then becomes possible, for example, to construct antibodies whose variable region has a higher specific affinity for a particular antigen, or whose constant region can elicit enhanced complement fixation, or to make other improvements in properties possessed by a particular constant region.

Another example is "altered antibodies", which refers to antibodies in which the naturally occurring amino acid sequence in a vertebrate antibody has been varies. Utilizing recombinant DNA techniques, antibodies can be redesigned to obtain desired characteristics. The possible variations are many, and range from the changing of one or more amino acids to the complete redesign of a region, for example, the constant region. Changes in the constant region, in general, to attain desired cellular process characteristics, e.g., changes in complement fixation, interaction with membranes, and other effector functions. Changes in the variable region can be made to alter antigen binding characteristics. The antibody can also be engineered to aid the specific delivery of a molecule or substance to a specific cell or tissue site. The desired alterations can be made by known techniques in molecular biology, e.g., recombinant techniques, site-directed mutagenesis, etc.

Yet another example are "univalent antibodies", which are aggregates comprised of a heavy-chain/light-chain dimer bound to the Fc (i.e., stem) region of a second heavy chain. This type of antibody escapes antigenic modulation. See, e.g., Glennie et al. Nature 295: 712 (1982). Included also within the definition of antibodies are "Fab" fragments of antibodies. The "Fab" region refers to those portions of the heavy and light chains which are roughly equivalent, or analogous, to the sequences which comprise the branch portion of the heavy and light chains, and which have been shown to exhibit immunological binding to a specified antigen, but which lack the effector Fc portion. "Fab" includes aggregates of one heavy and one light chain (commonly known as Fab'), as well as tetramers containing the 2H and 2L chains (referred to as F(ab)2), which are capable of selectively reacting with a designated antigen or antigen family. Fab antibodies can be divided into subsets analogous to those described above, i.e., "vertebrate Fab", "hybrid Fab", "chimeric Fab", and "altered Fab". Methods of producing Fab fragments of antibodies are known within the art and include, for example, proteolysis, and synthesis by recombinant techniques.

"Antigen-antibody complex" refers to the complex formed by an antibody that is specifically bound to an epitope on an antigen.

"Immunogenic polypeptide" refers to a polypeptide that elicits a cellular and/or humoral immune response in a mammal, whether alone or linked to a carrier, in the presence or absence of an adjuvant.

"Antigenic determinant" refers to the site on an antigen or hapten to which a specific antibody molecule or specific cell surface receptor binds.

As used herein, "treatment" refers to any of (i) the prevention of infection or reinfection, as in a traditional vaccine, (ii) the reduction or elimination of symptoms, and (iii) the substantial or complete elimination of the pathogen in question. Treatment may be effected prophylactically (prior to infection) or therapeutically (following infection).

By "vertebrate subject" is meant any member of the subphylum cordata, including, without limitation, humans and other primates, including non-human primates such as chimpanzees and other apes and monkey species; farm animals such as cattle, sheep, pigs, goats and horses; domestic mammals such as dogs and cats; laboratory animals including

rodents such as mice, rats and guinea pigs; birds, including domestic, wild and game birds such as chickens, turkeys and other gallinaceous birds, ducks, geese, and the like. The term does not denote a particular age. Thus, both adult and newborn individuals are intended to be covered. The invention described herein is intended for use in any of the above vertebrate species, since the immune systems of all of these vertebrates operate similarly.

II. Modes of Carrying out the Invention

Before describing the present invention in detail, it is to be understood that this invention is not limited to particular formulations or process parameters as such may, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments of the invention only, and is not intended to be limiting.

Although a number of compositions and methods similar or equivalent to those described herein can be used in the practice of the present invention, the preferred materials and methods are described herein.

General Overview

An aim of an HCV vaccine is to generate broad immunity to a wide breadth of antigens because HCV is so divergent and because humoral as well as cellular immune responses are desirable to combat this human pathogen. While antibodies generated against the envelope glycoprotein(s) might aid in virus neutralization, there is additional benefit to be derived from a vaccine that includes other regions. The likelihood of T-helper responses generated against a polypeptide would be helpful in a vaccine setting as would generation of cytotoxic T cells. The non-structural region represents such a candidate antigen, but processing by the protease generates several polypeptides, making purification complicated. It would be advantageous, therefore, to derive a non-structural cassette that is unprocessed by the NS3 protease.

The present invention solves this and other problems using compositions and methods involving an N-terminal deletion in NS3, which removes the catalytic domain.

As such, some or all of the remainder of the non-structural region (through NS5B) is expressed as an intact polypeptide. Expression of this species has been documented in mammalian cells as well as in yeast. Further, in certain aspects, polynucleotides encoding HCV core polypeptides (or fragments thereof) are added (*e.g,*. operably linked) to the carboxy-terminus of the non-structural cassette. As the core coding region is relatively highly conserved among HCV isolates, the presence of this region may enhance the immune response. Because core has at its C-terminus a very hydrophobic domain (amino acids 174-191), shorter versions of core were also engineered onto the polypeptide. As described in detail herein, the truncation of core to amino acid 121 yielded higher expression than the amino acid 173 truncation when engineered onto the C-terminus of the mutant NS polypeptide. The combination of most of the non-structural region fused to a C-terminally truncated core into a polypeptide is novel and has advantages for vaccine immunization. Moreover, because the aim is not necessarily to generate antibody responses to this polypeptide, there is no need to maintain a native conformation, enabling a more facile purification protocol.

Mutant HCV Non-Structural Polypeptides

Genomes of HCV strains contain a single open reading frame of approximately 9,000 to 12,000 nucleotides, which is transcribed into a polyprotein. An HCV polyprotein is cleaved to produce at least ten distinct products, in the order of NH_2- Core-E1-E2-p7-NS2-NS3-NS4a-NS4b-NS5a-NS5b-COOH. Mutant HCV polypeptides of the invention contain an N-terminal deletion in NS3, which removes or disables the catalytic domain. Preferably, the polypeptides also include the remainder of the non-structural region, although in certain embodiments, the polypeptides may include less than all of the remaining NS polypeptides, for example mutant NS polypeptides including any combinations of NS2-NS3-NS4a-NS4b-NS5a-NS5b (*e.g.,* NS3NS3-NS5a-NS5b; NS3-NS4a-NS4b; NS3-NS4a-NS4b-NS5a; NS3-NS4b-NS5a-NS5b; NS3-NS4a-NS5a; NS3-NS4b-NS5a; NS3-NS4b-NS5b; etc.).

The HCV NS3 protein functions as a protease and a helicase and occurs at approximately amino acid 1027 to amino acid 1657 of the polyprotein (numbered relative

to HCV-1). *See* Choo *et al.* (1991) Proc. Natl. Acad. Sci. USA 88:2451-2455. HCV NS4 occurs at approximately amino acid 1658 to amino acid 1972, NS5a occurs at approximately amino acid 1973 to amino acid 2420, and HCV NS5b occurs at approximately amino acid 2421 to amino acid 3011 of the polyprotein (numbered relative to HCV-1) (Choo *et al.*, 1991).

The mutant polypeptides described herein can either be full-length polypeptides or portions of NS3, NS4 (NS4a and NS4b), NS5a, and NS5b polypeptides. Epitopes of NS3, NS4 (NS4a and NS4b), NS5a, NS5b, NS3NS4NS5a, and NS3NS4NS5aNS5b can be identified by several methods. For example, NS3, NS4, NS5a, NS5b polypeptides or fusion proteins comprising any combination of the above, can be isolated, for example, by immunoaffinity purification using a monoclonal antibody for the polypeptide or protein. The isolated protein sequence can then be screened by preparing a series of short peptides by proteolytic cleavage of the purified protein, which together span the entire protein sequence. By starting with, for example, 100-mer polypeptides, each polypeptide can be tested for the presence of epitopes recognized by a T cell receptor on an HCV-activated T cell, progressively smaller and overlapping fragments can then be tested from an identified 100-mer to map the epitope of interest.

Epitopes recognized by a T cell receptor on an HCV-activated T cell can be identified by, for example, ^{51}Cr release assay (see Example 2) or by lymphoproliferation assay (see Example 4). In a ^{51}Cr release assay, target cells can be constructed that display the epitope of interest by cloning a polynucleotide encoding the epitope into an expression vector and transforming the expression vector into the target cells. Non-structural polypeptides can occur in any order in the fusion protein. If desired, at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more of one or more of the polypeptides may occur in the fusion protein. Multiple viral strains of HCV occur, and NS3, NS4, NS5a, and NS5b polypeptides of any of these strains can be used in a fusion protein.

Nucleic acid and amino acid sequences of a number of HCV strains and isolates, including nucleic acid and amino acid sequences of NS3, NS4, NS5a, NS5b genes and polypeptides have been determined. For example, isolate HCV J1.1 is described in Kubo *et al.* (1989) Japan. Nucl. Acids Res. 17:10367-10372; Takeuchi *et al.*(1990) Gene

91:287-291; Takeuchi *et al.* (1990) J. Gen. Virol. 71:3027-3033; and Takeuchi *et al.* (1990) Nucl. Acids Res. 18:4626. The complete coding sequences of two independent isolates, HCV-J and BK, are described by Kato *et al.*, (1990) Proc. Natl. Acad. Sci. USA 87:9524-9528 and Takamizawa *et al.*, (1991) J. Virol. 65:1105-1113 respectively.

Publications that describe HCV-1 isolates include Choo *et al.* (1990) Brit. Med. Bull. 46:423-441; Choo *et al.* (1991) Proc. Natl. Acad. Sci. USA 88:2451-2455 and Han *et al.* (1991) Proc. Natl. Acad. Sci. USA 88:1711-1715. HCV isolates HC-J1 and HC-J4 are described in Okamoto *et al.* (1991) Japan J. Exp. Med. 60:167-177. HCV isolates HCT 18~, HCT 23, Th, HCT 27, EC1 and EC10 are described in Weiner *et al.* (1991) Virol. 180:842-848. HCV isolates Pt-1, HCV-K1 and HCV-K2 are described in Enomoto *et al.* (1990) Biochem. Biophys. Res. Commun. 170:1021-1025. HCV isolates A, C, D & E are described in Tsukiyama-Kohara *et al.* (1991) Virus Genes 5:243-254.

Each of the mutant HCV polypeptides containing at least portions of NS3, NS4 and NS5 can be obtained from the same HCV strain or isolate or from different HCV strains or isolates. Thus, each non-structural region of the polypeptide can be from the same HCV strain or isolate or from each different HCV strains or isolates. In addition to the mutant HCV non-structural polypeptides described herein, the proteins can contain other polypeptides derived from the HCV polyprotein. For example, it may be desirable to include polypeptides derived from the core region of the HCV polyprotein. This region occurs at amino acid positions 1-191 of the HCV polyprotein, numbered relative to HCV-1. Either the full-length protein or epitopes of the full-length protein may be used in the subject fusions, such as those epitopes found between amino acids 10-53, amino acids 10-45, amino acids 67-88, amino acids 120-130, or any of the core epitopes identified in, e.g., Houghton et al., U.S. Patent No. 5,350,671; Chien et al., *Proc. Natl. Acad. Sci. USA* (1992) 89:10011-10015; Chien et al., *J. Gastroent. Hepatol.* (1993) 8:S33-39; Chien et al., International Publication No. WO 93/00365; Chien, D.Y., International Publication No. WO 94/01778; and commonly owned, U.S. Patent No. 6,150,087, the disclosures of which are incorporated herein by reference in their entireties. When present, additional non-structural HCV polypeptides such as core can be obtained from the same HCV strain or isolate or from different HCV strains or isolates.

Preferably, the above-described mutant proteins, as well as the individual components of these proteins, are produced recombinantly. A polynucleotide encoding these proteins can be introduced into an expression vector which can be expressed in a suitable expression system. A variety of bacterial, yeast, mammalian, insect and plant expression systems are available in the art and any such expression system can be used. Optionally, a polynucleotide encoding these proteins can be translated in a cell-free translation system. Such methods are well known in the art. The proteins also can be constructed by solid phase protein synthesis.

If desired, the mutant polypeptides, or the individual components of these polypeptides, also can contain other amino acid sequences, such as amino acid linkers or signal sequences, as well as ligands useful in protein purification, such as glutathione-S-transferase and staphylococcal protein A.

Polynucleotides

The polynucleotides of the present invention are not necessarily physically derived from the nucleotide sequences shown, but can be generated in any manner, including, for example, chemical synthesis or DNA replication or reverse transcription or transcription. In addition, combinations of regions corresponding to that of the designated sequences can be modified in ways known to the art to be consistent with an intended use.

The DNA encoding the desired polypeptide, whether in fused or mature form, and whether or not containing a signal sequence to permit secretion, can be ligated into expression vectors suitable for any convenient host. Both eukaryotic and prokaryotic host systems are presently used in forming recombinant polypeptides, and a summary of some of the more common control systems and host cell is given below. The polypeptide produced in such host cells is then isolated from lysed cells or from the culture medium and purified to the extent needed for its intended use.

Purification can be by techniques known in the art, for example, differential extraction, salt fractionation, chromatography on ion exchange resins, affinity chromatography, centrifugation, alkali resolubilization of insoluble protein, and the like. See, for example, Methods in Enzymology for a variety of methods for purifying proteins.

Polynucleotides contain less than an entire HCV genome and can be RNA or single- or double-stranded DNA. Preferably, the polynucleotides are isolated free of other components, such as proteins and lipids. Polynucleotides of the invention can also comprise other nucleotide sequences, such as sequences coding for linkers, signal sequences, or ligands useful in protein purification such as glutathione-S-transferase and staphylococcal protein A.

Polynucleotides encoding mutant HCV non-structural polypeptides can be isolated from a genomic library derived from nucleic acid sequences present in, for example, the plasma, serum, or liver homogenate of an HCV infected individual or can be synthesized in the laboratory, for example, using an automatic synthesizer. An amplification method such as PCR can be used to amplify polynucleotides from either HCV genomic DNA or cDNA.

Further, while the polypeptides that are not NS3, NS4, or NS5 of HCV of the present invention can comprise a substantially complete viral domain, in many applications all that is required is that the polypeptide comprise an antigenic or immunogenic region of the virus. An antigenic region of a polypeptide is generally relatively small-typically 8 to 10 amino acids or less in length. Fragments of as few as 5 amino acids can characterize an antigenic region. These segments can correspond to regions of, for example, C, E1, or E2 epitopes. Accordingly, using the cDNAs of C, E1, or E2 as a basis, DNAs encoding short segments of C, E1, or E2 polypeptides can be expressed recombinantly either as fusion proteins, or as isolated polypeptides. In addition, short amino acid sequences can be conveniently obtained by chemical synthesis.

Polynucleotides encoding the polypeptides described herein can comprise coding sequences for these polypeptides which occur naturally or can be artificial sequences which do not occur in nature. These polynucleotides can be ligated to form a coding sequence for the fusion proteins using standard molecular biology techniques. If desired, polynucleotides can be cloned into an expression vector and transformed into, for example, bacterial, yeast, insect, plant or mammalian cells so that the fusion proteins of the invention can be expressed in and isolated from a cell culture.

The expression of polypeptides containing these domains in a variety of recombinant host cells, including, for example, bacteria, yeast, insect, plant and vertebrate

cells, give rise to important immunological reagents which can be used for diagnosis, detection, and vaccines.

The general techniques used in extracting the genome from a virus, preparing and probing a cDNA library, sequencing clones, constructing expression vectors, transforming cells, performing immunological assays such as radioimmunoassays and. ELISA assays, for growing cells in culture, and the like are known in the art and laboratory manuals are available describing these techniques. However, as a general guide, the following sets forth some sources currently available for such procedures, and for materials useful in carrying them out.

Both prokaryotic and eukaryotic host cells may be used for expression of desired coding sequences when appropriate control sequences which are compatible with the designated host are used. Among prokaryotic hosts, E. coli is most frequently used. Expression control sequences for prokaryotes include promoters, optionally containing operator portions, and ribosome binding sites. Transfer vectors compatible with prokaryotic hosts are commonly derived from, for example, pBR322, a plasmid containing operons conferring ampicillin and tetracycline resistance, and the various pUC vectors, which also contain sequences conferring antibiotic resistance markers. These markers may be used to obtain successful transformants by selection. Commonly used prokaryotic control sequences include the Beta-lactamase (penicillinase) and lactose promoter systems (Chang et al. (1977), Nature 198:1056), the tryptophan (trp) promoter system (Goeddel et al. (1980) Nucleic Acid Res. 8:4057), the lambda-derived P[L]promoter and N gene ribosome binding site (Shimatake et al. (1981) Nature 292:128) and the hybrid tac promoter (De Boer et al. (1983) Proc. Natl. Acad. Sci. U.S.A. 292:128) derived from sequences of the trp and lac UV5 promoters. The foregoing systems are particularly compatible with E. coli; if desired, other prokaryotic hosts such as strains of Bacillus or Pseudomonas may be used, with corresponding control sequences.

Eukaryotic hosts include mammalian and yeast cells in culture systems. Mammalian cell lines available as hosts for expression are known in the art and include many immortalized cell lines available from the American Type Culture Collection (ATCC), including HeLa cells, Chinese hamster ovary (CHO) cells, baby hamster kidney

(BHK) cells, and a number of other cell lines. Suitable promoters for mammalian cells are also known in the art and include viral promoters such as that from Simian Virus 40 (SV40) (Fiers (1978), Nature 273:113), Rous sarcoma virus (RSV), adenovirus (ADV), and bovine papilloma virus (BPV). Mammalian cells may also require terminator sequences and poly A addition sequences; enhancer sequences which increase expression may also be included, and sequences which cause amplification of the gene may also be desirable. These sequences are known in the art. Vectors suitable for replication in mammalian cells may include viral replicons, or sequences which insure integration of the appropriate sequences encoding NANBV epitopes into the host genome.

The vaccinia virus system can also be used to express foreign DNA in mammalian cells. To express heterologous genes, the foreign DNA is usually inserted into the thymidine kinase gene of the vaccinia virus and then infected cells can be selected. This procedure is known in the art and further information can be found in these references (Mackett et al. J. Virol. 49: 857-864 (1984) and Chapter 7 in DNA Cloning, Vol. 2, IRL Press).

Yeast expression systems are also known to one of ordinary skill in the art. A yeast promoter is any DNA sequence capable of binding yeast RNA polymerase and initiating the downstream (3') transcription of a coding sequence (*e.g.*, structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site (the "TATA Box") and a transcription initiation site. A yeast promoter may also have a second domain called an upstream activator sequence (UAS), which, if present, is usually distal to the structural gene. The UAS permits regulated (inducible) expression. Constitutive expression occurs in the absence of a UAS. Regulated expression may be either positive or negative, thereby either enhancing or reducing transcription.

Yeast is a fermenting organism with an active metabolic pathway, therefore sequences encoding enzymes in the metabolic pathway provide particularly useful promoter sequences. Examples include alcohol dehydrogenase (ADH) (EP-A-0 284 044), enolase, glucokinase, glucose-6-phosphate isomerase, glyceraldehyde-3-phosphate-

dehydrogenase (GAP or GAPDH), hexokinase, phosphofructokinase, 3-phosphoglycerate mutase, and pyruvate kinase (PyK) (EPO-A-0 329 203). The yeast *PHO5* gene, encoding acid phosphatase, also provides useful promoter sequences (Myanohara *et al.* (1983) *Proc. Natl. Acad. Sci. USA 80*:1).

In addition, synthetic promoters which do not occur in nature also function as yeast promoters. For example, UAS sequences of one yeast promoter may be joined with the transcription activation region of another yeast promoter, creating a synthetic hybrid promoter. Examples of such hybrid promoters include the ADH regulatory sequence linked to the GAP transcription activation region (US Patent Nos. 4,876,197 and 4,880,734). Other examples of hybrid promoters include promoters which consist of the regulatory sequences of either the *ADH2*, *GAL4*, *GAL10*, OR *PHO5* genes, combined with the transcriptional activation region of a glycolytic enzyme gene such as GAP or PyK (EP-A-0 164 556). Furthermore, a yeast promoter can include naturally occurring promoters of non-yeast origin that have the ability to bind yeast RNA polymerase and initiate transcription. Examples of such promoters include, *inter alia*, (Cohen *et al.* (1980) *Proc. Natl. Acad. Sci. USA 77*:1078; Henikoff *et al.* (1981) *Nature 283*:835; Hollenberg *et al.* (1981) *Curr. Topics Microbiol. Immunol. 96*:119; Hollenberg *et al.* (1979) "The Expression of Bacterial Antibiotic Resistance Genes in the Yeast Saccharomyces cerevisiae," in: *Plasmids of Medical, Environmental and Commercial Importance* (eds. K.N. Timmis and A. Puhler); Mercerau-Puigalon *et al.* (1980) *Gene 11*:163; Panthier *et al.* (1980) *Curr. Genet. 2*:109).

A DNA molecule may be expressed intracellularly in yeast. A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus of the recombinant protein will always be a methionine, which is encoded by the ATG start codon. If desired, methionine at the N-terminus may be cleaved from the protein by *in vitro* incubation with cyanogen bromide.

Fusion proteins provide an alternative for yeast expression systems, as well as in mammalian, baculovirus, and bacterial expression systems. Usually, a DNA sequence encoding the N-terminal portion of an endogenous yeast protein, or other stable protein, is fused to the 5' end of heterologous coding sequences. Upon expression, this construct will

provide a fusion of the two amino acid sequences. For example, the yeast or human superoxide dismutase (SOD) gene, can be linked at the 5' terminus of a foreign gene and expressed in yeast. The DNA sequence at the junction of the two amino acid sequences may or may not encode a cleavable site. See *e.g.*, EP-A-0 196 056. Another example is a ubiquitin fusion protein. Such a fusion protein is made with the ubiquitin region that preferably retains a site for a processing enzyme (*e.g.*, ubiquitin-specific processing protease) to cleave the ubiquitin from the foreign protein. Through this method, therefore, native foreign protein can be isolated (*e.g.*, WO88/024066).

Alternatively, foreign proteins can also be secreted from the cell into the growth media by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provide for secretion in yeast of the foreign protein. Preferably, there are processing sites encoded between the leader fragment and the foreign gene that can be cleaved either *in vivo* or *in vitro*. The leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell.

DNA encoding suitable signal sequences can be derived from genes for secreted yeast proteins, such as the yeast invertase gene (EP-A-0 012 873; JPO. 62,096,086) and the A-factor gene (US patent 4,588,684). Alternatively, leaders of non-yeast origin, such as an interferon leader, exist that also provide for secretion in yeast (EP-A-0 060 057).

A preferred class of secretion leaders are those that employ a fragment of the yeast alpha-factor gene, which contains both a "pre" signal sequence, and a "pro" region. The types of alpha-factor fragments that can be employed include the full-length pre-pro alpha factor leader (about 83 amino acid residues) as well as truncated alpha-factor leaders (usually about 25 to about 50 amino acid residues) (US Patents 4,546,083 and 4,870,008; EP-A-0 324 274). Additional leaders employing an alpha-factor leader fragment that provides for secretion include hybrid alpha-factor leaders made with a presequence of a first yeast, but a pro-region from a second yeast alphafactor. (*e.g.*, see WO 89/02463.)

Usually, transcription termination sequences recognized by yeast are regulatory regions located 3' to the translation stop codon, and thus together with the promoter flank the coding sequence. These sequences direct the transcription of an mRNA which can be

translated into the polypeptide encoded by the DNA. Examples of transcription terminator sequence and other yeast-recognized termination sequences, such as those coding for glycolytic enzymes.

Usually, the above described components, comprising a promoter, leader (if desired), coding sequence of interest, and transcription termination sequence, are put together into expression constructs. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (*e.g.,* plasmids) capable of stable maintenance in a host, such as yeast or bacteria. The replicon may have two replication systems, thus allowing it to be maintained, for example, in yeast for expression and in a prokaryotic host for cloning and amplification. Examples of such yeast-bacteria shuttle vectors include YEp24 (Botstein *et al.* (1979) *Gene 8*:17-24), pCl/1 (Brake *et al.* (1984) *Proc. Natl. Acad. Sci USA 81*:4642-4646), and YRp17 (Stinchcomb *et al.* (1982) *J. Mol. Biol. 158*:157). In addition, a replicon may be either a high or low copy number plasmid. A high copy number plasmid will generally have a copy number ranging from about 5 to about 200, and usually about 10 to about 150. A host containing a high copy number plasmid will preferably have at least about 10, and more preferably at least about 20. Enter a high or low copy number vector may be selected, depending upon the effect of the vector and the foreign protein on the host. See *e.g.,* Brake *et al., supra.*

Alternatively, the expression constructs can be integrated into the yeast genome with an integrating vector. Integrating vectors usually contain at least one sequence homologous to a yeast chromosome that allows the vector to integrate, and preferably contain two homologous sequences flanking the expression construct. Integrations appear to result from recombinations between homologous DNA in the vector and the yeast chromosome (Orr-Weaver *et al.* (1983) *Methods in Enzymol. 101*:228-245). An integrating vector may be directed to a specific locus in yeast by selecting the appropriate homologous sequence for inclusion in the vector. See Orr-Weaver *et al., supra.* One or more expression construct may integrate, possibly affecting levels of recombinant protein produced (Rine *et al.* (1983) *Proc. Natl. Acad. Sci. USA 80*:6750). The chromosomal sequences included in the vector can occur either as a single segment in the vector, which results in the integration of the entire vector, or two segments homologous to adjacent

segments in the chromosome and flanking the expression construct in the vector, which can result in the stable integration of only the expression construct.

Usually, extrachromosomal and integrating expression constructs may contain selectable markers to allow for the selection of yeast strains that have been transformed. Selectable markers may include biosynthetic genes that can be expressed in the yeast host, such as *ADE2*, *HIS4*, *LEU2*, *TRP1*, and *ALG7*, and the G418 resistance gene, which confer resistance in yeast cells to tunicamycin and G418, respectively. In addition, a suitable selectable marker may also provide yeast with the ability to grow in the presence of toxic compounds, such as metal. For example, the presence of *CUP1* allows yeast to grow in the presence of copper ions (Butt *et al.* (1987) *Microbiol, Rev. 51*:351).

Alternatively, some of the above described components can be put together into transformation vectors. Transformation vectors are usually comprised of a selectable marker that is either maintained in a replicon or developed into an integrating vector, as described above.

Expression and transformation vectors, either extrachromosomal replicons or integrating vectors, have been developed for transformation into many yeasts. For example, expression vectors have been developed for, *inter alia*, the following yeasts: Candida albicans (Kurtz, *et al.* (1986) *Mol. Cell. Biol. 6*:142), Candida maltosa (Kunze, *et al.* (1985) *J. Basic Microbiol. 25*:141). Hansenula polymorpha (Gleeson, *et al.* (1986) *J. Gen. Microbiol. 132*:3459; Roggenkamp *et al.* (1986) *Mol. Gen. Genet. 202*:302), Kluyveromyces fragilis (Das, *et al.* (1984) *J. Bacteriol. 158*:1165), Kluyveromyces lactis (De Louvencourt *et al.* (1983) *J. Bacteriol. 154*:737; Van den Berg *et al.* (1990) *Bio/Technology 8*:135), Pichia guillerimondii (Kunze *et al.* (1985) *J. Basic Microbiol. 25*:141), Pichia pastoris (Cregg, *et al.* (1985) *Mol. Cell. Biol. 5*:3376; US Patent Nos. 4,837,148 and 4,929,555), Saccharomyces cerevisiae (Hinnen *et al.* (1978) *Proc. Natl. Acad. Sci. USA 75*:1929; Ito *et al.* (1983) *J. Bacteriol. 153*:163), Schizosaccharomyces pombe (Beach and Nurse (1981) *Nature 300*:706), and Yarrowia lipolytica (Davidow, *et al.* (1985) *Curr. Genet. 10*:380471 Gaillardin, *et al.* (1985) *Curr. Genet. 10*:49).

Methods of introducing exogenous DNA into yeast hosts are well-known in the art, and usually include either the transformation of spheroplasts or of intact yeast cells treated

with alkali cations. Transformation procedures usually vary with the yeast species to be transformed. (See *e.g.,* Kurtz *et al.* (1986) *Mol. Cell. Biol. 6*:142; Kunze *et al.* (1985) *J. Basic Microbiol. 25*:141; Candida; Gleeson *et al.* (1986) *J. Gen. Microbiol. 132*:3459; Roggenkamp *et al.* (1986) *Mol. Gen. Genet. 202*:302; Hansenula; Das *et al.* (1984) *J. Bacteriol. 158*:1165; De Louvencourt *et al.* (1983) *J. Bacteriol. 154*:1165; Van den Berg *et al.* (1990) *Bio/Technology 8*:135; Kluyveromyces; Cregg *et al.* (1985) *Mol. Cell. Biol. 5*:3376; Kunze *et al.* (1985) *J. Basic Microbiol. 25*:141; US Patent Nos. 4,837,148 and 4,929,555; Pichia; Hinnen *et al.* (1978) *Proc. Natl. Acad. Sci. USA 75*;1929; Ito *et al.* (1983) *J. Bacteriol. 153*:163 Saccharomyces; Beach and Nurse (1981) *Nature 300*:706; Schizosaccharomyces; Davidow *et al.* (1985) *Curr. Genet. 10*:39; Gaillardin *et al.* (1985) *Curr. Genet. 10*:49; Yarrowia).

Bacterial expression techniques are known in the art. A bacterial promoter is any DNA sequence capable of binding bacterial RNA polymerase and initiating the downstream (3') transcription of a coding sequence (*e.g.,* structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site and a transcription initiation site. A bacterial promoter may also have a second domain called an operator, that may overlap an adjacent RNA polymerase binding site at which RNA synthesis begins. The operator permits negative regulated (inducible) transcription, as a gene repressor protein may bind the operator and thereby inhibit transcription of a specific gene. Constitutive expression may occur in the absence of negative regulatory elements, such as the operator. In addition, positive regulation may be achieved by a gene activator protein binding sequence, which, if present is usually proximal (5') to the RNA polymerase binding sequence. An example of a gene activator protein is the catabolite activator protein (CAP), which helps initiate transcription of the lac operon in Escherichia coli (E. coli) (Raibaud *et al.* (1984) *Annu. Rev. Genet. 18*:173). Regulated expression may therefore be either positive or negative, thereby either enhancing or reducing transcription.

Expression and transformation vectors, either extra-chromosomal replicons or integrating vectors, have been developed for transformation into many bacteria. For

example, expression vectors have been developed for, *inter alia*, the following bacteria: Bacillus subtilis (Palva *et al.* (1982) *Proc. Natl. Acad. Sci. USA 79*:5582; EP-A-0 036 259 and EP-A-0 063 953; WO 84/04541), Escherichia coli (Shimatake *et al.* (1981) *Nature 292*:128; Amann *et al.* (1985) *Gene 40*:183; Studier *et al.* (1986) *J. Mol. Biol. 189*:113; EP-A-0 036 776,EP-A-0 136 829 and EP-A-0 136 907), Streptococcus cremoris (Powell *et al.* (1988) *Appl. Environ. Microbiol. 54*:655); Streptococcus lividans (Powell *et al.* (1988) *Appl. Environ. Microbiol. 54*:655), Streptomyces lividans (US patent 4,745,056).

Methods of introducing exogenous DNA into bacterial hosts are well-known in the art, and usually include either the transformation of bacteria treated with $CaCl_2$ or other agents, such as divalent cations and DMSO. DNA can also be introduced into bacterial cells by electroporation. Transformation procedures usually vary with the bacterial species to be transformed. (See *e.g.*, Masson *et al.* (1989) *FEMS Microbiol. Lett. 60*:273; Palva *et al.* (1982) *Proc. Natl. Acad. Sci. USA 79*:5582; EP-A-0 036 259 and EP-A-0 063 953; WO 84/04541, Bacillus, Miller *et al.* (1988) *Proc. Natl. Acad. Sci. 85*:856; Wang *et al.* (1990) *J. Bacteriol. 172*:949; Campylobacter, Cohen *et al.* (1973) *Proc. Natl. Acad. Sci. 69*:2110; Dower *et al.* (1988) *Nucleic Acids Res. 16*:6127; Kushner (1978) "An improved method for transformation of Escherichia coli with ColE1-derived plasmids. *In Genetic Engineering: Proceedings of the International Symposium on Genetic Engineering* (eds. H.W. Boyer and S. Nicosia); Mandel *et al.* (1970) *J. Mol. Biol. 53*:159; Taketo (1988) *Biochim. Biophys. Acta 949*:318; Escherichia; Chassy *et al.* (1987) *FEMS Microbiol. Lett. 44*:173 Lactobacillus; Fiedler *et al.* (1988) *Anal. Biochem 170*:38, Pseudomonas; Augustin *et al.* (1990) *FEMS Microbiol. Lett. 66*:203, Staphylococcus, Barany *et al.* (1980) *J. Bacteriol. 144*:698; Harlander (1987) "Transformation of Streptococcus lactis by electroporation, in: *Streptococcal Genetics* (ed. J. Ferretti and R. Curtiss III); Perry *et al.* (1981) *Infect. Immun. 32*:1295; Powell *et al.* (1988) *Appl. Environ. Microbiol. 54*:655; Somkuti *et al.* (1987) *Proc. 4th Evr. Cong. Biotechnology 1*:412, Streptococcus).

In addition, viral antigens can be expressed in insect cells by the Baculovirus system. A general guide to Baculovirus expression by Summer and Smith is A Manual of

Methods for Baculovirus Vectors and Insect Cell Culture Procedures (Texas Agricultural Experiment Station Bulletin No. 1555). To incorporate the heterologous gene into the Baculovirus genome the gene is first cloned into a transfer vector containing some Baculovirus sequences. This transfer vector, when it is cotransfected with wild-type virus into insect cells, will recombine with the wild-type virus. Usually, the transfer vector will be engineered so that the heterologous gene will disrupt the wild-type Baculovirus polyhedron gene. This disruption enables easy selection of the recombinant virus since the cells infected with the recombinant virus will appear phenotypically different from the cells infected with the wild-type virus. The purified recombinant virus can be used to infect cells to express the heterologous gene. The foreign protein can be secreted into the medium if a signal peptide is linked in frame to the heterologous gene; otherwise, the protein will be bound in the cell lysates. For further information, see Smith et al Mol. & Cell. Biol. 3:2156-2165 (1983) or Luckow and Summers in Virology 17: 31-39 (1989).

Baculovirus expression can also be affected in plant cells. There are many plant cell culture and whole plant genetic expression systems known in the art. Exemplary plant cellular genetic expression systems include those described in patents, such as: US 5,693,506; US 5,659,122; and US 5,608,143. Additional examples of genetic expression in plant cell culture has been described by Zenk, *Phytochemistry* 30:3861-3863 (1991). Descriptions of plant protein signal peptides may be found in addition to the references described above in Vaulcombe et al., *Mol. Gen. Genet.* 209:33-40 (1987); Chandler et al., *Plant Molecular Biology* 3:407-418 (1984); Rogers, *J. Biol. Chem.* 260:3731-3738 (1985); Rothstein et al., *Gene* 55:353-356 (1987); Whittier et al., Nucleic Acids Research 15:2515-2535 (1987); Wirsel et al., *Molecular Microbiology* 3:3-14 (1989); Yu et al., *Gene* 122:247-253 (1992). A description of the regulation of plant gene expression by the phytohormone, gibberellic acid and secreted enzymes induced by gibberellic acid can be found in R.L. Jones and J. MacMillin, Gibberellins: in: *Advanced Plant Physiology*,. Malcolm B. Wilkins, ed., 1984 Pitman Publishing Limited, London, pp. 21-52. References that describe other metabolically-regulated genes: Sheen, *Plant Cell*, 2:1027-1038(1990); Maas et al., *EMBO J.* 9:3447-3452 (1990); Benkel and Hickey, *Proc. Natl. Acad. Sci.* 84:1337-1339 (1987).

All plants from which protoplasts can be isolated and cultured to give whole regenerated plants can be transformed by the present invention so that whole plants are recovered which contain the transferred gene. It is known that practically all plants can be regenerated from cultured cells or tissues, including but not limited to all major species of sugarcane, sugar beet, cotton, fruit and other trees, legumes and vegetables. Some suitable plants include, for example, species from the genera *Fragaria*, *Lotus*, *Medicago*, *Onobrychis*, *Trifolium*, *Trigonella*, *Vigna*, *Citrus*, *Linum*, *Geranium*, *Manihot*, *Daucus*, *Arabidopsis*, *Brassica*, *Raphanus*, *Sinapis*, *Atropa*, *Capsicum*, *Datura*, *Hyoscyamus*, *Lycopersion*, *Nicotiana*, *Solanum*, *Petunia*, *Digitalis*, *Majorana*, *Cichorium*, *Helianthus*, *Lactuca*, *Bromus*, *Asparagus*, *Antirrhinum*, *Hererocallis*, *Nemesia*, *Pelargonium*, *Panicum*, *Pennisetum*, *Ranunculus*, *Senecio*, *Salpiglossis*, *Cucumis*, *Browaalia*, *Glycine*, *Lolium*, *Zea*, *Triticum*, *Sorghum*, and *Datura*.

Transformation can be by any method for introducing polynucleotides into a host cell, including, for example packaging the polynucleotide in a virus and transducing a host cell with the virus, and by direct uptake of the polynucleotide. The transformation procedure used depends upon the host to be transformed. Bacterial transformation by direct uptake generally employs treatment with calcium or rubidium chloride (Cohen (1972), Proc. Natl. Acad. Sci. U.S.A. 69:2110; Maniatis et al. (1982), MOLECULAR CLONING; A LABORATORY MANUAL (Cold Spring Harbor Press, Cold Spring Harbor, N.Y.). Yeast transformation by direct uptake may be carried out using the method of Hinnen et al. (1978) Proc. Natl. Acad. Sci. U.S.A. 75: 1929. Mammalian transformations by direct uptake may be conducted using the calcium phosphate precipitation method of Graham and Van der Eb (1978), Virology 52:546 or the various known modifications thereof.

Vector construction employs techniques which are known in the art. Site-specific DNA cleavage is performed by treating with suitable restriction enzymes under conditions which generally are specified by the manufacturer of these commercially available enzymes. The cleaved fragments may be separated using polyacrylamide or agarose gel electrophoresis techniques, according to the general procedures found in Methods in Enzymology (1980) 65:499-560. Sticky ended cleavage fragments may be blunt ended using E. coli DNA polymerase I (Klenow) in the presence of the appropriate

deoxynucleotide triphosphates (dNTPs) present in the mixture. Treatment with S1 nuclease may also be used, resulting in the hydrolysis of any single stranded DNA portions.

Ligations are carried out using standard buffer and temperature conditions using T4 DNA ligase and ATP; sticky end ligations require less ATP and less ligase than blunt end ligations. When vector fragments are used as part of a ligation mixture, the vector fragment is often treated with bacterial alkaline phosphatase (BAP) or calf intestinal alkaline phosphatase to remove the 5'-phosphate and thus prevent religation of the vector; alternatively, restriction enzyme digestion of unwanted fragments can be used to prevent ligation. Ligation mixtures are transformed into suitable cloning hosts, such as E. coli, and successful transformants selected by, for example, antibiotic resistance, and screened for the correct construction.

Synthetic oligonucleotides may be prepared using an automated oligonucleotide synthesizer as described by Warner (1984), DNA 3:401. If desired, the synthetic strands may be labeled with ^{32}P by treatment with polynucleotide kinase in the presence of ^{32}P-ATP, using standard conditions for the reaction. DNA sequences, including those isolated from cDNA libraries, may be modified by known techniques, including, for example site directed mutagenesis, as described by Zoller (1982), Nucleic Acids Res. 10:6487.

The expression constructs of the present invention, including the desired fusion, or individual expression constructs comprising the individual components of these fusions, may be used for nucleic acid immunization, to activate HCV-specific T cells, using standard gene delivery protocols. Methods for gene delivery are known in the art. See, e.g., U.S. Patent Nos. 5,399,346, 5,580,859, 5,589,466, incorporated by reference herein in their entireties. Genes can be delivered either directly to the vertebrate subject or, alternatively, delivered *ex vivo*, to cells derived from the subject and the cells reimplanted in the subject. For example, the constructs can be delivered as plasmid DNA, e.g., contained within a plasmid, such as pBR322, pUC, or ColE1

Additionally, the expression constructs can be packaged in liposomes prior to delivery to the cells. Lipid encapsulation is generally accomplished using liposomes which are able to stably bind or entrap and retain nucleic acid. The ratio of condensed DNA to lipid preparation can vary but will generally be around 1:1 (mg DNA:micromoles lipid), or

more of lipid. For a review of the use of liposomes as carriers for delivery of nucleic acids, see, Hug and Sleight, *Biochim. Biophys. Acta.* (1991) 1097:1-17; Straubinger et al., in *Methods of Enzymology* (1983), Vol. 101, pp. 512-527.

Liposomal preparations for use with the present invention include cationic (positively charged), anionic (negatively charged) and neutral preparations, with cationic liposomes particularly preferred. Cationic liposomes are readily available. For example, N[1-2,3-dioleyloxy)propyl]-N,N,N-triethylammonium (DOTMA) liposomes are available under the trademark Lipofectin, from GIBCO BRL, Grand Island, NY. (See, also, Felgner et al., *Proc. Natl. Acad. Sci. USA* (1987) 84:7413-7416). Other commercially available lipids include transfectace (DDAB/DOPE) and DOTAP/DOPE (Boerhinger). Other cationic liposomes can be prepared from readily available materials using techniques well known in the art. See, e.g., Szoka et al., *Proc. Natl. Acad. Sci. USA* (1978) 75:4194-4198; PCT Publication No. WO 90/11092 for a description of the synthesis of DOTAP (1,2-bis(oleoyloxy)-3-(trimethylammonio)propane) liposomes. The various liposome-nucleic acid complexes are prepared using methods known in the art. See, e.g., Straubinger et al., in METHODS OF IMMUNOLOGY (1983), Vol. 101, pp. 512-527; Szoka et al., *Proc. Natl. Acad. Sci. USA* (1978) 75:4194-4198; Papahadjopoulos et al., *Biochim. Biophys. Acta* (1975) 394:483; Wilson et al., *Cell* (1979) 17:77); Deamer and Bangham, *Biochim. Biophys. Acta* (1976) 443:629; Ostro et al., *Biochem. Biophys. Res. Commun.* (1977) 76:836; Fraley et al., *Proc. Natl. Acad. Sci. USA* (1979) 76:3348); Enoch and Strittmatter, *Proc. Natl. Acad. Sci. USA* (1979) 76:145); Fraley et al., *J. Biol. Chem.* (1980) 255:10431; Szoka and Papahadjopoulos, *Proc. Natl. Acad. Sci. USA* (1978) 75:145; and Schaefer-Ridder et al., *Science* (1982) 215:166.

The DNA can also be delivered in cochleate lipid compositions similar to those described by Papahadjopoulos et al., *Biochem. Biophys. Acta.* (1975) 394:483-491. See, also, U.S. Patent Nos. 4,663,161 and 4,871,488.

A number of viral based systems have been developed for gene transfer into mammalian cells. For example, retroviruses provide a convenient platform for gene delivery systems, such as murine sarcoma virus, mouse mammary tumor virus, Moloney murine leukemia virus, and Rous sarcoma virus. A selected gene can be inserted into a

vector and packaged in retroviral particles using techniques known in the art. The recombinant virus can then be isolated and delivered to cells of the subject either *in vivo* or *ex vivo*. A number of retroviral systems have been described (U.S. Patent No. 5,219,740; Miller and Rosman, *BioTechniques* (1989) 7:980-990; Miller, A.D., *Human Gene Therapy* (1990) 1:5-14; Scarpa et al., *Virology* (1991) 180:849-852; Burns et al., *Proc. Natl. Acad. Sci. USA* (1993) 90:8033-8037; and Boris-Lawrie and Temin, *Cur. Opin. Genet. Develop.* (1993) 3:102-109. Briefly, retroviral gene delivery vehicles of the present invention may be readily constructed from a wide variety of retroviruses, including for example, B, C, and D type retroviruses as well as spumaviruses and lentiviruses such as FIV, HIV, HIV-1, HIV-2 and SIV (see RNA Tumor Viruses, Second Edition, Cold Spring Harbor Laboratory, 1985). Such retroviruses may be readily obtained from depositories or collections such as the American Type Culture Collection ("ATCC"; 10801 University Blvd., Manassas, VA 20110-2209), or isolated from known sources using commonly available techniques.

A number of adenovirus vectors have also been described, such as adenovirus Type 2 and Type 5 vectors. Unlike retroviruses which integrate into the host genome, adenoviruses persist extrachromosomally thus minimizing the risks associated with insertional mutagenesis (Haj-Ahmad and Graham, *J. Virol.* (1986) 57:267-274; Bett et al., *J. Virol.* (1993) 67:5911-5921; Mittereder et al., *Human Gene Therapy* (1994) 5:717-729; Seth et al., *J. Virol.* (1994) 68:933-940; Barr et al., *Gene Therapy* (1994) 1:51-58; Berkner, K.L. *BioTechniques* (1988) 6:616-629; and Rich et al., *Human Gene Therapy* (1993) 4:461-476).

Molecular conjugate vectors, such as the adenovirus chimeric vectors described in Michael et al., *J. Biol. Chem.* (1993) 268:6866-6869 and Wagner et al., *Proc. Natl. Acad. Sci. USA* (1992) 89:6099-6103, can also be used for gene delivery.

Members of the Alphavirus genus, such as but not limited to vectors derived from the Sindbis and Semliki Forest viruses, VEE, will also find use as viral vectors for delivering the gene of interest. For a description of Sindbis-virus derived vectors useful for the practice of the instant methods, see, Dubensky et al., *J. Virol.* (1996) 70:508-519; and International Publication Nos. WO 95/07995 and WO 96/17072.

e.g., U.S. Patent Nos. 4,945,050; 5,036,006; 5,100,792; 5,179,022; 5,371,015; and 5,478,744.

Compositions

The invention also provides compositions comprising the HCV polypeptides or polynucleotides described herein. Such compositions are useful as diagnostics, for example, using the mutant polypeptides (or polynucleotides encoding these polypeptides) in diagnostic reagents. Diagnostics using polypeptides and polynucleotides are known to those of skill in the art.

In addition, immunogenic compounds can be prepared from one or more immunogenic polypeptides derived from the polypeptides described herein, for example the ΔNS35 polypeptide. The preparation of immunogenic compounds which contain immunogenic polypeptide(s) as active ingredients is known to one skilled in the art. Typically, such immunogenic compounds are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid prior to injection can also be prepared. The preparation can also be emulsified, or the protein encapsulated in liposomes.

Immunogenic and diagnostic compositions of the invention preferably comprise a pharmaceutically acceptable carrier. The carrier should not itself induce the production of antibodies harmful to the host. Pharmaceutically acceptable carriers are well known to those in the art. Such carriers include, but are not limited to, large, slowly metabolized, macromolecules, such as proteins, polysaccharides such as latex functionalized sepharose, agarose, cellulose, cellulose beads and the like, polylactic acids, polyglycolic acids, polymeric amino acids such as polyglutamic acid, polylysine, and the like, amino acid copolymers, and inactive virus particles.

Pharmaceutically acceptable salts can also be used in compositions of the invention, for example, mineral salts such as hydrochlorides, hydrobromides, phosphates, or sulfates, as well as salts of organic acids such as acetates, proprionates, malonates, or benzoates. Especially useful protein substrates are serum albumins, keyhole limpet hemocyanin, immunoglobulin molecules, thyroglobulin, ovalbumin, tetanus toxoid, and

Other vectors can be used, including but not limited to simian virus 40, cytomegalovirus. Bacterial vectors, such as Salmonella ssp. *Yersinia enterocolitica*, Shigella spp., *Vibrio cholerae*, Mycobacterium strain BCG, and *Listeria monocytogenes* can be used. Minichromosomes such as MC and MC1, bacteriophages, cosmids (plasmids into which phage lambda *cos* sites have been inserted) and replicons (genetic elements that are capable of replication under their own control in a cell) can also be used.

The expression constructs may also be encapsulated, adsorbed to, or associated with, particulate carriers. Such carriers present multiple copies of a selected molecule to the immune system and promote trapping and retention of molecules in local lymph nodes. The particles can be phagocytosed by macrophages and can enhance antigen presentation through cytokine release. Examples of particulate carriers include those derived from polymethyl methacrylate polymers, as well as microparticles derived from poly(lactides) and poly(lactide-co-glycolides), known as PLG. See, e.g., Jeffery et al., *Pharm. Res.* (1993) 10:362-368; and McGee et al., *J. Microencap.* (1996).

A wide variety of other methods can be used to deliver the expression constructs to cells. Such methods include DEAE dextran-mediated transfection, calcium phosphate precipitation, polylysine- or polyornithine-mediated transfection, or precipitation using other insoluble inorganic salts, such as strontium phosphate, aluminum silicates including bentonite and kaolin, chromic oxide, magnesium silicate, talc, and the like. Other useful methods of transfection include electroporation, sonoporation, protoplast fusion, liposomes, peptoid delivery, or microinjection. See, e.g., Sambrook et al., *supra*, for a discussion of techniques for transforming cells of interest; and Felgner, P.L., *Advanced Drug Delivery Reviews* (1990) 5:163-187, for a review of delivery systems useful for gene transfer. One particularly effective method of delivering DNA using electroporation is described in International Publication No. WO/0045823.

Additionally, biolistic delivery systems employing particulate carriers such as gold and tungsten, are especially useful for delivering the expression constructs of the present invention. The particles are coated with the construct to be delivered and accelerated to high velocity, generally under a reduced atmosphere, using a gun powder discharge from a "gene gun." For a description of such techniques, and apparatuses useful therefore, see,

other proteins well known to those of skill in the art. Compositions of the invention can also contain liquids or excipients, such as water, saline, glycerol, dextrose, ethanol, or the like, singly or in combination, as well as substances such as wetting agents, emulsifying agents, or pH buffering agents. Liposomes can also be used as a carrier for a composition of the invention, such liposomes are described above.

If desired, co-stimulatory molecules which improve immunogen presentation to lymphocytes, such as B7-1 or B7-2, or cytokines such as GM-CSF, IL-2, and IL-12, can be included in a composition of the invention. Optionally, adjuvants can also be included in a composition. Adjuvants which can be used include, but are not limited to: (1) aluminum salts (alum), such as aluminum hydroxide, aluminum phosphate, aluminum sulfate, etc; (2) oil-in-water emulsion formulations (with or without other specific immunostimulating agents such as muramyl peptides (see below) or bacterial cell wall components), such as for example (a) MF59 (PCT Publ. No. WO 90/14837), containing 5% Squalene, 0.5% Tween 80, and 0.5% Span 85 (optionally containing various amounts of MTP-PE), formulated into submicron particles using a microfluidizer such as Model 110Y microfluidizer (Microfluidics, Newton, MA), (b) SAF, containing 10% Squalane, 0.4% Tween 80, 5% pluronic-blocked polymer L121, and thr-MDP (see below) either microfluidized into a submicron emulsion or vortexed to generate a larger particle size emulsion, and (c) Ribi™ adjuvant system (RAS), (Ribi Immunochem, Hamilton, MT) containing 2% Squalene, 0.2% Tween 80, and one or more bacterial cell wall components from the group consisting of monophosphorylipid A (MPL), trehalose dimycolate (TDM), and cell wall skeleton (CWS), preferably MPL + CWS (Detox™); (3) saponin adjuvants, such as Stimulon™ (Cambridge Bioscience, Worcester, MA) may be used or particles generated therefrom such as ISCOMs (immunostimulating complexes); (4) Complete Freund's Adjuvant (CFA) and Incomplete Freund's Adjuvant (IFA); (5) cytokines, such as interleukins (e.g., IL-1, IL-2, IL-4, IL-5, IL-6, IL-7, IL-12, *etc.*), interferons (e.g., gamma interferon), macrophage colony stimulating factor (M-CSF), tumor necrosis factor (TNF), etc; (6) detoxified mutants of a bacterial ADP-ribosylating toxin such as a cholera toxin (CT), a pertussis toxin (PT), or an E. coli heat-labile toxin (LT), particularly LT-K63, LT-R72, CT-S109, PT-K9/G129; see, e.g., WO 93/13302 and WO 92/19265; (7) other

substances that act as immunostimulating agents to enhance the effectiveness of the composition; and (8) microparticles with adsorbed macromolecules, as described in copending U.S. Patent Application Serial No. 09/285,855 (filed April 2, 1999) and international Patent Application Serial No. PCT/US99/17308 (filed July 29, 1999). Alum and MF59 are preferred. The effectiveness of an adjuvant can be determined by measuring the amount of antibodies directed against an immunogenic polypeptide containing an HCV antigenic sequence resulting from administration of this polypeptide in immunogenic compounds which are also comprised of the various adjuvants.

As mentioned above, muramyl peptides include, but are not limited to, N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), –acetyl-normuramyl-L-alanyl-D-isoglutamine (CGP 11637, referred to nor-MDP), N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-*sn*-glycero-3-hydroxyphosphoryloxy)-ethylamine (CGP 19835A, referred to as MTP-PE), *etc.*

Thus, such recombinant or synthetic HCV polypeptides can be used in vaccines and as diagnostics. Further, antibodies raised against these polypeptides can also be used as diagnostics, or for passive immunotherapy. In addition, antibodies to these polypeptides are useful for isolating and identifying HCV particles.

Native HCV antigens can also be isolated from HCV virions. The virions can be grown in HCV infected cells in tissue culture, or in an infected host.

Administration and Delivery

The polynucleotide and polypeptide compositions described herein (*e.g.*, immunogenic compounds) may be administered to a subject using any suitable delivery means. Methods of delivering nucleic acids into host cells are discussed above. Further, HCV polynucleotides and/or polypeptides can be administered parenterally, by injection, usually, subcutaneously, intramuscularly, transdermally or transcutaneously. Certain adjuvants, e.g. LTK63, LTR72 or PLG formulations, can be administered intranasally or orally. Additional formulations which are suitable for other modes of administration include suppositories. For suppositories, traditional binders and carriers can include, for example, polyalkylene glycols or triglycerides; such suppositories can be formed from

mixtures containing the active ingredient in the range of 0.5% to 10%, preferably 1%-2%. Other oral formulations include such normally employed excipients as, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, and the like. These compositions take the form of solutions, suspensions, tablets, pills, capsules, sustained release formulations or powders and contain 10%-95% of active ingredient, preferably 25%-70%.

The polypeptides of the present invention can be formulated into the immunogenic compound as neutral or salt forms. Pharmaceutically acceptable salts include the acid addition salts (formed with free amino groups of the peptide) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids such as acetic, oxalic, tartaric, maleic, and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, 2-ethylamino ethanol, histidine, procaine, and the like.

The immunogenic compounds are administered in a manner compatible with the dosage formulation, and in such amount as will be prophylactically and/or therapeutically effective. The quantity to be administered, which is generally in the range of 5 micrograms to 250 micrograms of polypeptide per dose, depends on the subject to be treated, capacity of the subject's immune system to synthesize antibodies, and the degree of protection desired. Precise amounts of active ingredient required to be administered may depend on the judgment of the practitioner and can be peculiar to each subject.

The immunogenic compound can be given in a single dose schedule, or preferably in a multiple dose schedule. A multiple dose schedule is one in which a primary course of vaccination can be with 1-10 separate doses, followed by other doses given at subsequent time intervals required to maintain and or reenforce the immune response, for example, at 1-4 months for a second dose, and if needed, a subsequent dose(s) after several months. Further, the course of administration may include polynucleotides and polypeptides, together or sequentially (for example, priming with a polynucleotide composition and boosting with a polypeptide composition). The dosage regimen will also, at least in part,

be determined by the need of the individual and be dependent upon the judgment of the practitioner.

In certain embodiments, administration of the polynucleotides and polypeptides described herein is used to activate T cells. In addition to the practical advantages of simplicity of construction and modification, administration of polynucleotides encoding mutant NS polypeptides results in the synthesis of a mutant NS polypeptide in the host. Thus, these immunogens are presented to the host immune system with native post-translational modifications, structure, and conformation. The polynucleotides are preferably injected intramuscularly to a large mammal, such as a human, at a dose of 0.5, 0.75, 1.0, 1.5, 2.0, 2.5, 5 or 10 mg/kg.

The proteins and/or polynucleotides can be administered either to a mammal which is not infected with an HCV or can be administered to an HCV-infected mammal. The particular dosages of the polynucleotides or fusion proteins in a composition or will depend on many factors including, but not limited to the species, age, and general condition of the mammal to which the composition is administered, and the mode of administration of the composition. An effective amount of the composition of the invention can be readily determined using only routine experimentation. *In vitro* and *in vivo* models can be employed to identify appropriate doses. Generally, 0.5, 0.75, 1.0, 1.5, 2.0, 2.5, 5 or 10 mg will be administered to a large mammal, such as a baboon, chimpanzee, or human. If desired, co-stimulatory molecules or adjuvants can also be provided before, after, or together with the compositions.

Antibodies and Diagnostics

Antibodies, both monoclonal and polyclonal, which are directed against HCV epitopes are particularly useful in diagnosis, and those which are neutralizing are useful in passive immunotherapy. Monoclonal antibodies, in particular, may be used to raise anti-idiotype antibodies.

Anti-idiotype antibodies are immunoglobulins which carry an "internal image" of the antigen of the infectious agent against which protection is desired. Techniques for raising anti-idiotype antibodies are known in the art. See, e.g., Grzych (1985), Nature

316:74; MacNamara et al. (1984), Science 226:1325, Uytdehaag et al (1985), J. Immunol. 134:1225. These anti-idiotype antibodies may also be useful for treatment and/or diagnosis of NANBH, as well as for an elucidation of the immunogenic regions of HCV antigens.

An immunoassay for viral antigen may use, for example, a monoclonal antibody directed towards a viral epitope, a combination of monoclonal antibodies directed towards epitopes of one viral polypeptide, monoclonal antibodies directed towards epitopes of different viral polypeptides, polyclonal antibodies directed towards the same viral antigen, polyclonal antibodies directed towards different viral antigens or a combination of monoclonal and polyclonal antibodies.

Immunoassay protocols may be based, for example, upon competition, or direct reaction, or sandwich type assays. Protocols may also, for example, use solid supports, or may be by immunoprecipitation. Most assays involve the use of labeled antibody or polypeptide. The labels may be, for example, fluorescent, chemiluminescent, radioactive, or dye molecules. Assays which amplify the signals from the probe are also known. Examples of which are assays which utilize biotin and avidin, and enzyme-labeled and mediated immunoassays, such as ELISA assays.

An enzyme-linked immunosorbent assay (ELISA) can be used to measure either antigen or antibody concentrations. This method depends upon conjugation of an enzyme to either an antigen or an antibody, and uses the bound enzyme activity as a quantitative label. To measure antibody, the known antigen is fixed to a solid phase (e.g., a microplate or plastic cup), incubated with test serum dilutions, washed, incubated with anti-immunoglobulin labeled with an enzyme, and washed again. Enzymes suitable for labeling are known in the art, and include, for example, horseradish peroxidase. Enzyme activity bound to the solid phase is measured by adding the specific substrate, and determining product formation or substrate utilization colorimetrically. The enzyme activity bound is a direct function of the amount of antibody bound.

To measure antigen, a known specific antibody is fixed to the solid phase, the test material containing antigen is added, after an incubation the solid phase is washed, and a second enzyme-labeled antibody is added. After washing, substrate is added, and enzyme activity is estimated colorimetrically, and related to antigen concentration.

The HCV fusion proteins, such as NS3 mutant and core fusion proteins, can also be used to produce HCV-specific polyclonal and monoclonal antibodies. HCV-specific polyclonal and monoclonal antibodies specifically bind to HCV antigens.

Polyclonal antibodies can be produced by administering the fusion protein to a mammal, such as a mouse, a rabbit, a goat, or a horse. Serum from the immunized animal is collected and the antibodies are purified from the plasma by, for example, precipitation with ammonium sulfate, followed by chromatography, preferably affinity chromatography. Techniques for producing and processing polyclonal antisera are known in the art.

Monoclonal antibodies directed against HCV-specific epitopes present in the fusion proteins can also be readily produced. Normal B cells from a mammal, such as a mouse, immunized with, e.g., a mutant NS3 polypeptide or NS-core fusion protein can be fused with, for example, HAT-sensitive mouse myeloma cells to produce hybridomas. Hybridomas producing HCV-specific antibodies can be identified using RIA or ELISA and isolated by cloning in semi-solid agar or by limiting dilution. Clones producing HCV-specific antibodies are isolated by another round of screening.

Antibodies, either monoclonal and polyclonal, which are directed against HCV epitopes, are particularly useful for detecting the presence of HCV or HCV antigens in a sample, such as a serum sample from an HCV-infected human. An immunoassay for an HCV antigen may utilize one antibody or several antibodies. An immunoassay for an HCV antigen may use, for example, a monoclonal antibody directed towards an HCV epitope, a combination of monoclonal antibodies directed towards epitopes of one HCV polypeptide, monoclonal antibodies directed towards epitopes of different HCV polypeptides, polyclonal antibodies directed towards the same HCV antigen, polyclonal antibodies directed towards different HCV antigens, or a combination of monoclonal and polyclonal antibodies. Immunoassay protocols may be based, for example, upon competition, direct reaction, or sandwich type assays using, for example, labeled antibody. The labels may be, for example, fluorescent, chemiluminescent, or radioactive.

The polyclonal or monoclonal antibodies may further be used to isolate HCV particles or antigens by immunoaffinity columns. The antibodies can be affixed to a solid support by, for example, adsorption or by covalent linkage so that the antibodies retain

their immunoselective activity. Optionally, spacer groups may be included so that the antigen binding site of the antibody remains accessible. The immobilized antibodies can then be used to bind HCV particles or antigens from a biological sample, such as blood or plasma. The bound HCV particles or antigens are recovered from the column matrix by, for example, a change in pH.

Methods of Eliciting Immune Responses

HCV-specific T cells that are activated by the above-described polypeptides, expressed *in vivo* or *in vitro* preferably recognize an epitope of an HCV polypeptide such as a mutant NS3 polypeptide, including an epitope of a mutant HCV polypeptide. HCV-specific T cells can be $CD8^+$ or $CD4^+$.

HCV-specific $CD8^+$ T cells preferably are cytotoxic T lymphocytes (CTL) which can kill HCV-infected cells that display NS3, NS4, NS5a, NS5b epitopes complexed with an MHC class I molecule. HCV-specific $CD8^+$ T cells may also express interferon-γ (IFN-γ). HCV-specific $CD8^+$ T cells can be detected by, for example, ^{51}Cr release assays. ^{51}Cr release assays measure the ability of HCV-specific $CD8^+$ T cells to lyse target cells displaying an nonstructural (*e.g.*, mutant NS) epitope. HCV-specific $CD8^+$ T cells which express IFN-γ can also be detected by immunological methods, preferably by intracellular staining for IFN-γ after *in vitro* stimulation with a mutant NS polypeptide.

HCV-specific $CD4^+$ cells activated by the above-described polypeptides, expressed *in vivo* or *in vitro*, and combinations of the individual components of these proteins, preferably recognize an epitope of a mutant non-structural polypeptide, including an epitope of a mutant protein, that is bound to an MHC class II molecule on an HCV-infected cell and proliferate in response to stimulating mutant peptides.

HCV-specific $CD4^+$ T cells can be detected by a lymphoproliferation assay. Lymphoproliferation assays measure the ability of HCV-specific $CD4^+$ T cells to proliferate in response to an epitope.

Mutant NS (or fusions thereof with core, envelope or other viral polypeptides) can be used to activate HCV-specific T cells either *in vitro* or *in vivo*. Activation of HCV-specific T cells can be used, *inter alia*, to provide model systems to optimize CTL

responses to HCV and to provide prophylactic or therapeutic treatment against HCV infection. For *in vitro* activation, proteins are preferably supplied to T cells via a plasmid or a viral vector, such as an adenovirus vector, as described above.

Polyclonal populations of T cells can be derived from the blood, and preferably from peripheral lymphoid organs, such as lymph nodes, spleen, or thymus, of mammals that have been infected with an HCV. Preferred mammals include mice, chimpanzees, baboons, and humans. The HCV serves to expand the number of activated HCV-specific T cells in the mammal. The HCV-specific T cells derived from the mammal can then be restimulated *in vitro* by adding HCV epitopic peptides to the T cells. The HCV-specific T cells can then be tested for, *inter alia*, proliferation (*e.g.*,. lymphoproliferation assays known in the art), the production of IFN-γ, and the ability to lyse target cells displaying HCV NS epitopes *in vitro*.

The following examples are meant to illustrate the invention and are not meant to limit it in any way. Those of ordinary skill in the art will recognize modifications within the spirit and scope of the invention as set forth herein.

EXAMPLES

Example 1: Constructs

pCMV-II: pCMV-II (Figure 7, SEQ ID NO:5) was created to contain the human CMV promoter, enhancer, intron A, polylinker and the bovine growth hormone terminator in a deleted-pUC backbone (Life Technologies).

pT7-HCV: pT7-HCV was created in a polylinker-modified pUC vector to contain full-length HCV cDNA preceded by a synthetic T7 promoter. pT7-HCV also contains the complete 5' UTR and the poly A version of the 3' UTR.

pCMV.ΔNS35: To generate pCMV.ΔNS35 (Figure 5, SEQ ID NO:3), a two step procedure was undertaken. First, a PCR product was generated from pT7-HCV that corresponded to the following: a 5' EcoRI site, followed by the Kozak sequence of ACCATGG; the initiator ATG followed by amino acid #1242 and continuing to the StuI site. Second, the StuI to XbaI fragment from a full-length genomic clone was isolated. The genomic clone consisted of the T7 promoter fused to the full-length HCV cDNA with the poly A version of the 3' end, in a pUC vector. Finally, the EcoRI-StuI and StuI-XbaI fragments were ligated into the pCMV-II expression vector, transformed into HB101 competent cells and plated onto ampicillin (100 μg/ml). Miniprep analyses led to the identification of the desired clone which was amplified on a larger scale using a Quigen Gigaprep kit following the manufacturer's specifications. The resulting clone was named pCMV.ΔNS35 (Figure 5, SEQ ID NO:3).

pd.ΔNS3NS5: As shown schematically in Figure 10, the yeast expression plasmid pd.ΔNS3NS5 (SEQ ID NO:8) was constructed using restriction fragments obtained from the mammalian expression plasmid pCMV.KM.ΔNS35. pCMV.KM.ΔNS35 is identical to pCMV.ΔNS35 (Figure 5, SEQ ID NO:3) except that it contains a kanamycin resistance gene in the viral backbone. pCMV.KM.ΔNS35 was digested with EcoRI and NheI to obtain 2895bp EcoRI-NheI fragment. EcoRI-NheI fragment was ligated into pRSET HindIII-NheI subcloning vector with oligos (HE) from HindIII to EcoRI. After sequence verification, pRSETHindIII-NheI #6 was digested with HindIII and NheI to obtain a

2908bp HindIII-NheI fragment.

pCMV.KM.ΔNS35 was linearized with XbaI and ligated with synthetic oligos (XS) from XbaI-SalI. The ligation was digested with NheI and SalI to obtain 2481bp NheI-SalI fragment. The fragment was ligated into pET3a NheI-SalI subcloning vector. After sequence verification, pET3a NheI-SalI #2 was digested with NheI and SalI to obtain a 2481bp NheI-SalI fragment. BamHI-HindIII ADH2/GAPDH promoter fragment was then ligated with HindIII-NheI and NheI-SalI fragments into pBS24.1 BamHI-SalI yeast expression vector.

pd.ΔNS3NS5.PJ: pd.ΔNS3NS5.PJ (Figures 13 and 14; SEQ ID NO:10) was generated to create a "perfect junction" at the 5' and 3' end of the HCV coding region. At the 5' end of pd.ΔNS3NS5, there were 6 extra bases between the yeast ADH2/GAPDH promoter and the ATG of the polypeptide. At the 3' end, there were 52 bases of untranslated sequence between the stop codon of the polypeptide and the α-factor terminator in the yeast expression vector. pd.ΔNS3NS5.PJ was created by digesting pd.ΔNS3NS5 #17 with ScaI and SphI to obtain 4963bp ScaI-SphI fragment. pd.NS5b3011 was digested with SphI and SalI to obtain a 321bp SphI-SalI fragment which gave the "perfect junction" at the 3' end of the polypeptide. The ScaI-SphI and SphI-SalI fragments were ligated into pSP72 HindIII-SalI subcloning vector with synthetic oligos from HindIII-ScaI(HS) for the "perfect junction" at the 5' end.

The region of synthetic sequence in pSP72 HindIII-SalI clone# 6 was verified. pSP72 HindIII-SalI clone#6 was digested with HindIII and BlnI or with BlnI and SalI to obtain 2441bp HindIII-BlnI and 2895bp BlnI-SalI fragments, respectively. The BamHI-HindIII ADH2/GAPDH promoter fragment was ligated to HindIII-BlnI and BlnI-SalI fragments into pBS24.1 BamHI-SalI yeast expression vector.

pd.ΔNS3NS5.PJ.core121RT and pd.Δ NS3NS5.PJ.core173RT were generated and encode HCV core aa 1-121 at the C-terminus of the ΔNS3NS5 polypeptide (designated pd.ΔNS3NS5.PJ.core121RT, SEQ ID NO:12) and core aa 1-173 at the C-terminus of the ΔNS3NS5 polypeptide (designated pd.ΔNS3NS5.PJ.core173RT, SEQ ID NO:14). The core sequence had aa 9 mutated from Lys to Arg and aa 11 mutated from Asn to Thr,

designated as core 121RT or 173RT.

pd.ΔNS3NS5.PJ.core121RT and pd.ΔNS3NS5.PJ.core173RT: To generate pd.ΔNS3NS5.PJ.core121RT (Figure 17, SEQ ID NO:12) and pd.ΔNS3NS5.PJ.core173RT (Figure 18, SEQ ID NO:14). As shown in Figure 16, a NotI-Sal HCVcore121RT and HCVcore173RT were amplified by PCR, from an *E. coli* expression plasmid, pSODCF2.HCVcore191RT #2. Either the core 121RT Not-SalI PCR product or the core 173RT Not-SalI PCR product were ligated into a pT7Blue2 PstI-SalI subcloning vector with synthetic oligos (PN) from PstI to NotI. After sequence confirmation, pT7Blue2core121RT clone#9 and pT7Blue2core173RT clone#11 was digested with PstI and SalI to obtain 403bp and 559bp PstI-SalI fragments, respectively, for further cloning.

A 121bp NotI-PstI fragment from pSP72 HindIII-SalI clone #6 was isolated as described above during the cloning of pd.ΔNS3NS5.PJ. NotI-PstI and PstI-SalI fragments were assembled into a vector made by digesting pd.NS3NS5.PJ clone#5 (described above) with NotI and SalI.

ΔNS3NS5 and Core 140 and Core 150: An HCV core epitope was found which elicits CTLs in baboons (HCV core aa 121-135). Since pd.ΔNS3NS5.PJ.core121RT ends right before this potentially important epitope and was expressed better than the longer pd.ΔNS3NS5.PJ.core173RT construct (Example 2), two intermediate constructs were made which include this epitope, possibly giving intermediate expression levels. The two new constructs fused HCV core aa 1-140 or HCV core aa1-150 to the C terminus of ΔNS3NS5.PJ.

pd.ΔNS3NS5.PJ.core140RT (Figure 21, SEQ ID NO:16) and pd.ΔNS3NS5.PJ.core150RT (Figure 22, SEQ ID NO:18): As shown in Figure 20, a PstI-SalI HCVcore140RT and a PstI-SalIHCVcore150RT fragment were amplified by PCR from pd.ΔNS3NS5.PJ.core173RT clone #16. Ligate either HCV core PstI-SalI PCR products into pT7Blue2 PstI-SalI subcloning vector. After sequence confirmation, pT7Blue2core140RT clone#22 and pT7Blue2core150RT clone#26 were digested with PstI-SalI to obtain 460bp and 490bp PstI-SalI fragments, respectively, for further cloning.

A 121bp NotI-PstI fragment was isolated from pSP72 HindIII-SalI clone #6 (as described above during the cloning of pd.ΔNS3NS5.PJ. NotI-PstI and PstI-SalI fragments were assembled into a vector made by digesting pd.ΔNS3NS5.PJ clone#5 (described above) with NotI and SalI.

Example 2: Protein Expression

Various of the constructs described herein, encoding HCV-1 ΔNS3 to NS5 antigen (aa 1242-3011), were expressed in yeast. *S. cerevisiae* strain AD3 was transformed with pd.ΔNS3NS5 and checked for expression. A stained protein band at the expected molecular weight of 194 kD was not observed (Figure 12). Strain AD3 was also transformed with pd.ΔNS3NS5.PJ clone #5 and checked for expression. A protein band of the expected molecular weight of 194kD was detected (Figure 15). Strain AD3 was transformed with pd.ΔNS3NS5.PJ.core121RT clone #6 and pd.ΔNS3NS5.PJ.core173RT clone#15 and checked for expression. Protein bands of the expected molecular weight of 206kD and 210kD, respectively, were observed. Expression levels of the pd.ΔNS3NS5.PJ.core173RT construct were much less than that of the pd.ΔNS3NS5.PJ.core121RT construct. (See Figure19). Thus, there is a correlation of protein expression levels and the length of HCV core.

Strain AD3 were transformed with pd.ΔNS3NS5.PJ.core140RT clone# 29 and pd.ΔNS3NS5.PJ.core150RT clone#35 and checked for expression. Bands of the expected molecular weights of 208kD and 209kD were seen by stain at levels close to those of pd.ΔNS3NS5core173RT (Figure 23).

Example 3: Eliciting Immune Responses

A. Immunization

To evaluate the immunogenicity of the mutant NS polypeptides, studies using guinea pigs, rabbits, mice, rhesus macaques and/or baboons are performed. The studies are structured as follows: DNA immunization alone (single or multiple); DNA immunization followed by protein immunization (boost); DNA immunization followed by protein

immunization; immunization by PLG particles. Immunization is intramuscular or mucosally.

B. Humoral Immune Response

The humoral immune response is checked in serum specimens from immunized animals with anti-NS antibody ELISAs (enzyme-linked immunosorbent assays) at various times post-immunization. Briefly, serum from immunized animals is screened for antibodies directed against the NS or mutant NS proteins. Wells of ELISA microtiter plates are coated overnight with the selected HCV protein and washed four times; subsequently, blocking is done with PBS-0.2% Tween (Sigma). After removal of the blocking solution, diluted mouse serum is added. Sera are tested at various dilutions. Microtiter plates are washed and incubated with a secondary, peroxidase-coupled anti-mouse IgG antibody (Pierce, Rockford, IL). ELISA plates are washed and 3, 3', 5, 5'-tetramethyl benzidine (TMB; Pierce) is added per well. The optical density of each well is measured. Titers are typically reported as the reciprocal of the dilution of serum that gave a half-maximum optical density (O.D.). Similarly, generation of neutralization of binding (NOB) antibodies can be measured by methods known in the art.

C. Cellular Immune Response

The frequency of specific cytotoxic T-lymphocytes (CTL) is evaluated by a standard chromium release assay of peptide pulsed Balb/c mouse CD4 cells. Briefly, spleen cells (Effector cells, E) are obtained from the BALB/c mice immunized, cultured, restimulated, and assayed for CTL activity against HCV peptide-pulsed target cells. Cytotoxic activity is measured in a standard ^{51}Cr release assay.

Example 4: Immunization with PLG-delivered DNA.

The polylactide-co-glycolide (PLG) polymers are obtained from Boehringer Ingelheim, U.S.A. The PLG polymer is RG505, which has a copolymer ratio of 50/50 and a molecular weight of 65 kDa (manufacturers data). Cationic microparticles with adsorbed DNA are prepared using a modified solvent evaporation process, essentially as described in

Singh et al., *Proc. Natl. Acad. Sci. USA* (2000) 97:811-816. Briefly, the microparticles are prepared by emulsifying a 5% w/v polymer solution in methylene chloride with PBS at high speed using an IKA homogenizer. The primary emulsion is then added to distilled water containing cetyl trimethyl ammonium bromide (CTAB) (0.5% w/v). This results in the formation of a w/o/w emulsion which was stirred at room temperature, allowing the methylene chloride to evaporate. The resulting microparticles are washed in distilled water by centrifugation and freeze dried. Following preparation, washing and collection, DNA is adsorbed onto the microparticles by incubating cationic microparticles in a solution of DNA. The microparticles are then separated by centrifugation, the pellet washed with TE buffer and the microparticles are freeze dried, resuspended and administered to animals. Antibody titers are measured by ELISA assays.

All patents, patent applications, and other publications mentioned herein, are hereby incorporated herein by reference in their entireties.